



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number 155980

TO: Nita M Minnifield
Art Unit: 1645
Location: rem/3C01/3C18
Serial Number: 09/970076

Friday, June 17, 2005

From: Beverly Shears
Location: Biotech-Chem Library
REM 1A54
Phone: 571-272-2528
beverly.shears@uspto.gov

Search Notes

Protein Sequence Searches – February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension .rup) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (uniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

Revised
MM

THIS PAGE BLANK (USPTO)

XX 13-JUN-2002.
 PD 03-OCT-2001; 2001MO-US030941.
 XX 05-DEC-2000; 2000US-0251481P.
 PR (WISC) WISCONSIN ALUMNI RES FOUND.
 PA Young JAT, Bradley KA, Collier RJ, Mogridge JS;
 PI WPI; 2002-713235/77.
 DR N-PSDB; ABV73881.
 XX Novel isolated polypeptide useful for identifying agent that prevents or
 PT reduces effect of anthrax toxin on host cell, for treating human or non-
 PT human animal suffering from anthrax.
 XX
 PS Claim 1; Page 29-30; 45pp; English.
 CC The present sequence is the protein sequence of a human surface-bound
 CC anthrax toxin receptor (ATR), as predicted from an isolated cDNA clone.
 CC Anthrax toxin protective antigen (PA) binds to the ATR at a von
 CC Willebrand factor A domain located in the extracellular domain of ATR.
 CC The invention provides ATR polypeptides and polynucleotides, vectors,
 CC host cells, and transgenic and knock-out animals. It also provides
 CC methods for identifying molecules that bind the ATR and which reduce the
 CC toxicity of anthrax toxin. A claimed method for treating anthrax in a
 CC human or animal involves administering an agent that inhibits binding
 CC between PA and ATR at a level effective to reduce the severity of
 CC anthrax. Suitable agents include ATR or a PA-binding fragment of ATR, a
 CC PA-binding polypeptide at least 80% identical to these, a fusion protein,
 CC a monoclonal or polyclonal antibody, a polysaccharide, a lipid or a
 CC nucleic acid
 XX
 SQ Sequence 368 AA;
 Query Match 100.0%; Score 1536; DB 5; Length 368;
 Best Local Similarity 100.0%; Pred. No. 1.5e-156;
 Matches 295; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGRRRDEGPGACGGFDLYITLDKSGSVLHNNIYFVEQLAHKPSIPOLRMSFYVS 60
 DB 27 GGGRRRDEGPGACGGFDLYITLDKSGSVLHNNIYFVEQLAHKPSIPOLRMSFYVS 86
 QY 61 TRGTTMLKLTEDRQIRQGLBELQVLPFGSDTYMHGEPERASQIYYENRGYRTASYII 120
 DB 87 TRGTTMLKLTEDRQIRQGLBELQVLPFGSDTYMHGEPERASQIYYENRGYRTASYII 146
 QY 121 ALTDGELHEDLFFYSERERANRSRDGAIVYCVGVDFNETQLARIADSKDHVPVNDGFQ 180
 DB 147 ALTDGELHEDLFFYSERERANRSRDGAIVYCVGVDFNETQLARIADSKDHVPVNDGFQ 206
 QY 181 ALQGIHSLKSKSCIEIIAEPSTTCAGESQVYVVRNGRPHANNVDRVLCSEKINDSVT 240
 DB 207 ALQGIHSLKSKSCIEIIAEPSTTCAGESQVYVVRNGRPHANNVDRVLCSEKINDSVT 266
 QY 241 LNEKPESEVDYTLCPAPILKEVGKALQVSMNDGSLFISSVITTTTHCSDS 295
 DB 267 LNEKPESEVDYTLCPAPILKEVGKALQVSMNDGSLFISSVITTTTHCSDS 321
 RESULT 2
 AAE01439 standard; protein; 403 AA.
 AC AAE01439;
 DT 17-JUL-2001 (first entry)
 XX Human gene 4 encoded secreted protein HMLFR02, SEQ ID NO:94.
 XX Human; secreted protein; proliferative disorder; cancer; tumour;

KW foetal abnormality; developmental abnormality; haematopoietic disorder;
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
 KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
 KW cardiovascular disorder; angiosenic disorder; kidney disorder;
 KW gastrointestinal disorder; pregnancy-related disorder;
 KW endocrine disorder; infection; wound healing; vulnery; cell culture;
 KW chemotaxis; food additive; gene therapy; binding partner identification;
 KW chromosome 19.
 OS Homo sapiens.
 XX
 PH Key Location/Qualifiers
 FT Peptide 1..27
 FT /label= Signal_peptide
 FT Protein 28..403
 FT /note= "Mature human secreted protein"
 XX
 PN WO200134626-A1.
 XX
 PD 17-MAY-2001.
 XX
 PF 01-NOV-2000; 2000MO-US030045.
 XX
 PR 05-NOV-1999; 99US-0163581P.
 PR 30-JUN-2000; 2000US-0215133P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Ruben SM, Komatsoulis GA, Moore PA, Birse CE, Ni J;
 XX WPI; 2001-308778/32.
 DR N-PSDB; AAD05303.
 XX
 PT New nucleic acid molecules encoding 28 human secreted proteins for
 PT diagnosing, preventing, treating or ameliorating medical conditions and
 PT used as food additives or preservatives.
 PS Claim 11; Page 485-486; 562pp; English.
 XX
 CC AAD05300-AAD05319 represent cDNAs corresponding to 28 human secreted
 CC protein genes, and AAE01436-AAE01513 represent the proteins they encode.
 CC AAE01514-AAE01544 represent human secreted protein fragments or variants.
 CC The genes and their secreted proteins are useful for preventing, treating
 CC or ameliorating medical conditions, e.g., by protein or gene therapy.
 CC Pathological conditions can be diagnosed by determining the amount of the
 CC new protein in a sample or by determining the presence of mutations in
 CC the new genes. Specific uses are described for each of the 28 genes,
 CC based on the tissues in which they are most highly expressed, and include
 CC developing products for the diagnosis or treatment of proliferative
 CC disorders, cancer, tumours, foetal and developmental abnormalities,
 CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
 CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,
 CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
 CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
 CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
 CC angiogenic disorders, kidney disorders, gastrointestinal disorders,
 CC pregnancy-related disorders, endocrine disorders, and infections. The
 CC proteins can also be used to aid wound healing and epithelial cell
 CC proliferation, to prevent skin aging due to sunburn, to maintain organs
 CC before transplantation, for supporting cell culture of primary tissues,
 CC to regenerate tissues, to identify their cognate ligands or binding
 CC partners, and in chemotaxis, and can be used as a food additive or
 CC preservative to modify storage properties. Antibodies specific for a
 CC protein of the invention can be used in alleviating symptoms associated
 CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,
 CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The
 CC present sequence represents a human secreted protein of the invention
 XX
 SQ Sequence 403 AA;
 Query Match 100.0%; Score 1536; DB 4; Length 403;

	Best Local Similarity	100.0%	Pred. No. 1.8e-156	
Matches	295	Conservative	0	Mismatches 0; Indels 0; Gaps 0
QY	1	GGGGRRDDGGPACVGGFDLYFLIDKSGSVLHMMNEIYYFVEQLAHKFIISPQLRMSFIYFS	60	
Db	27	GGGGRRDDGGPACVGGFDLYFLIDKSGSVLHMMNEIYYFVEQLAHKFIISPQLRMSFIYFS	86	
QY	61	TRGTTMLMUTLDRDEIRIQQGLELQVLPFGDYYMHGEPBRASEQIYYENRQGYRTASVYI	120	
Db	87	TRGTTMLMUTLDRDEIRIQQGLELQVLPFGDYYMHGEPBRASEQIYYENRQGYRTASVYI	146	
QY	121	ALTTGELHEDLFFYSEREANRSDGAIYYCYGVDPFNETOLARIADSKDHYFPVNDGQ	180	
Db	147	ALTGELHEDLFFYSEREANRSDGAIYYCYGVDPFNETOLARIADSKDHYFPVNDGQ	206	
QY	181	ALQGIHSHILKKSCEIILAEPSTI CAGESFQVVVVRNGGFRHARNVDRVLCSEFKINDSVT	240	
Db	207	ALQGIHSHILKKSCEIILAEPSTI CAGESFQVVVVRNGGFRHARNVDRVLCSEFKINDSVT	266	
QY	241	LNKRPFSVEDPYLLCPAPILKEVGKKALQVSMNDGSLFISSSVIITTHHCGDGS	295	
Db	267	LNKRPFSVEDPYLLCPAPILKEVGKKALQVSMNDGSLFISSSVIITTHHCGDGS	321	

RESULT 3	
ABG63874	
ID	ABG63874 standard; protein; 403 AA.
XX	
AC	ABG63874;
XX	
DT	27-AUG-2002 (first entry)
XX	
XX	Human albumin fusion protein #549.
XX	

KM Albumin fusion protein; therapeutic protein X human albumin; HA
 KM human serum albumin; HSA; cancer; reproductive disorder;
 KM digestive disorder; immune disorder; endocrine disorder;
 KM hemetopietic disorder; neural disorder; connective disorder;
 KM cytotoxic; antifertility; antinflammatory; anticler;
 KM immunomodulator; anti-HIV; antidiabetic; hemostatic; nootropic;
 KM neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
 KM osteopathic; antitarrhetic.

OS Homo sapiens.
OS Synthetic.

PN WO200177137-A1.

PD 18-OCT-2001.

PF 12-APR-2001; 2001WO-US011988.

PR 12-APR-2000; 2000US-0229358P.

PR 21-DEC-2000; 2000US-0256931P.

PA (HUMA-) HUMAN GENOME SCI INC.
XX

P1 ROSEN CA, HABELLINE WA;
 XX

REF: 000070-2003
XX

albumin fused to a

PS Claim 1; Page 874-875; 2102pp; English.

CC The present invention relates to albumin

CC albumin HSA). The proteins are useful for treating a disease or disorder
CC that may be modulated by therapeutic protein X. The albumin extends the
CC shelf-life of protein X, and may increase its biological in vitro/in vivo
CC activity. The protein is useful for treating and diagnosing disorders

CC such as cancer, reproductive disorders, digestive disorders (e.g. Crohn's
CC disease, ulcerative colitis), immune disorders (e.g. acquired
CC immunodeficiency syndrome, AIDS), endocrine disorders (e.g. diabetes),
CC hematopoietic disorders, neural disorders (e.g. Alzheimer's,
CC Parkinson's, Creutzfeldt-Jacob disease, encephalomyelitis, meningitis,
CC schizophrenia), and connective disorders (e.g. osteoporosis, arthritis).
CC AB65326-AB65518 represent albumin fusion proteins of the invention
XX
Sequence 403 AA;
90

Query Match	100.0%;	Score 1536;	DB 5;	Length 403;
Best Local Similarity	100.0%;	Pred. No. 1.8e-156;		
Matches 295;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

RESULT 4
ADL77139
ID ADL77139 standard; protein; 403 AA

KM albumin:uric acid protein; cyrostatic; antinaeamic; antiarthritic;
KM antipneumatic; anti-HIV; immunosuppressive; antiinflammatory;
KM antiparasitic; antibacterial; osteopathic; dermatological; antigout;
KM immunomodulator; antiarrhythmic; cardiac; nootropic; antilipemic;
KM nephroprotective; uropathic; neuroprotective; antiparkinsonian; tranquilizer
KM antidiabetic; anabolic; hypertensive; vulnereary; gene therapy; cancer;
KM reproductive system disorder; therapeutic protein.

OS Unidentified.

PN US2004010134-A1

PD 15-JAN-2004

12-APR-2001; 2001US-00833245-
 XX

PK 12-APR-2000; 2000US-02293358P.
PB 25-APR-2000; 2000US-0199384P

[illegible]

PA (HASE/) HASELTINE W A.

PI Rosen CA, Habeltine V

DR WPI; 2004-090519/09.

PT New albumin fusion prote

PT or ameliorating diseases or disorders e.g. cancer, anemia, arthritis,
PT aschma, inflammatory bowel disease or Alzheimer's disease.
PS Disclosure; SEQ ID NO 621; 279pp; English.
XX
CC The invention relates to a novel albumin fusion protein. The invention
CC further relates to: a composition comprising the albumin fusion protein
CC and a pharmaceutical carrier; a kit comprising the composition of the
CC albumin fusion protein formula; a method of treating a disease or
CC disorder in a patient comprising the step of administering the albumin
CC fusion protein; a method of treating a patient with a disease or disorder
CC that is modulated by Therapeutic protein: X, or its fragment or variant;
CC a method of extending the shelf life of Therapeutic protein: X, or its
CC fragment or variant; a nucleic acid molecule comprising a polynucleotide
CC sequence encoding the albumin fusion protein; a vector comprising the
CC nucleic acid molecule of the albumin fusion protein; and a host cell
CC comprising the nucleic acid molecule of the albumin fusion protein. The
CC albumin fusion protein and its compositions have the following
CC activities: cytostatic, antiinfective, antiarthritic, antisthmatic, anti-
CC HIV, immunosuppressive, antiinflammatory, antiproliferative, antibacterial,
CC osteoplastic, dermatological, antipain, immunomodulator, antiaerhythmic,
CC cardiac, neurotropic, antiparkinsonian, tranquilizer, antidiabetic,
CC neuroprotective, antiparkinsonian, antiproliferative, anabolic,
CC hypertensive, and vulnerary. The albumin fusion protein nucleic acid may
CC be used in gene therapy to treat disorders. The albumin fusion protein is
CC useful for diagnosing, treating, preventing or ameliorating diseases or
CC disorders comprising indication: Y. The diseases or disorders include:
CC cancer (e.g. leukaemia, colon, bone, breast, liver or lung cancer),
CC immune or haematopoietic diseases (e.g. anaemia, Hodgkin's disease, acute
CC lymphocytic anaemia, multiple myeloma, arthritis, ashma, AIDS,
CC autoimmune disease, inflammatory bowel disease, psoriasis or Lyme
CC disease), reproductive system disorders (e.g. prostaticitis, inguinal
CC hernia, varicocele, penile carcinoma, ovarian adenocarcinoma or Sertoli-
CC Leydig tumours), musculoskeletal diseases (e.g. giant cell tumours,
CC Paget's disease, systemic lupus erythematosus, gout, muscular dystrophy
CC or cachexia), cardiovascular disease (e.g. rhabdomyomas, heart disease,
CC arrhythmia, cardiac arrest, heat valve disease, hypernatraemia or
CC hyponatraemia), mixed foetal diseases (e.g. foetal alcohol syndrome,
CC Down's syndrome, Patou syndrome, Turner's syndrome, Apert syndrome, or Tay-
CC Sachs disease), excretory diseases (e.g. urinary incontinence, urinary
CC tract infections or renal disorders), neural or sensory disease (e.g.
CC Alzheimer's disease, Parkinson's disease, cerebral malaria, meningitis,
CC cerebellar ataxia, attention deficit disorder, autism or obsessive
CC compulsive disorder), respiratory disease (e.g. emphysema, lung cancer or
CC occupational lung disease), endocrine diseases (e.g. diabetes, Addison's
CC disease or glomerulonephritis), digestive diseases (e.g. portal
CC hypertension, irritable bowel disease, gastric atrophy or pancreatitis)
CC or connective tissue or epithelial diseases (e.g. Crohn's disease,
CC scleroderma, wound healing or epidermolysis bullosa). This sequence
CC represents a therapeutic protein X relating to the albumin fusion protein
CC of the invention. The sequence listing data for this specification was
CC downloaded from the USPTO website.
XX
XX Sequence 403 AA;
SQ
Query Match 100.0%; Score 1536; DB 8; Length 403;
Best Local Similarity 100.0%; Pred. No. 1.8e-156;
Matches 295; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGGRRDGGPACVGGFDLYFLDKSGSVLHNMNIYFYEQLAHKFISPOLRMSFIYFS 60
DB 27 GGGGRRDGGPACVGGFDLYFLDKSGSVLHNMNIYFYEQLAHKFISPOLRMSFIYFS 86
QY 61 TRGTTMLKLTEDREIQIGLEBELQKVLPGDVTYHBEGERASBOIYYENROGYRTASVII 120
DB 87 TRGTTMLKLTEDREIQIGLEBELQKVLPGDVTYHBEGERASBOIYYENROGYRTASVII 146
QY 121 ALTDELHEDLFFYSERERANSRDIGAIYVCVGVDFNETOLARIADSKVHFVPYNDGFO 180
DB 147 ALTDELHEDLFFYSERERANSRDIGAIYVCVGVDFNETOLARIADSKVHFVPYNDGFO 206
QY 181 ALGQTHSLIKKSCIEILAAEPSTICAGSFQVVRGNGFRHARVNDVLSFKINDSVT 240

DB 207 ALGQTHSLIKKSCIEILAAEPSTICAGSFQVVRGNGFRHARVNDVLSFKINDSVT 266
QY 241 LNEKPSVEDTYLLCPAPILKEVGMKALQVSNNDGSLFISSVITTTTHCSDGS 295
DB 267 LNEKPSVEDTYLLCPAPILKEVGMKALQVSNNDGSLFISSVITTTTHCSDGS 321
RESULT 5
ABB90750
ID ABB90750 standard; protein; 564 AA.
XX
AC ABB90750;
XX
DT 30-MAY-2002 (first entry)
XX
XX Human Tumour Endothelial Marker polypeptide SEQ ID NO 232.
DE
XX Human; mouse; rat; TEM; tumour endothelial marker; NEM; PEM; cytostatic;
KW normal endothelial marker; pan-endothelial marker; immunostimulant;
KW antiangiogenic; tumour; neovascularization; vascularised tumour;
KW polycystic kidney disease; diabetes; retinopathy; rheumatoid arthritis;
KW psoriasis.
OS
XX Homo sapiens.
XX
XX W0200210217-A2.
XX
PD 07-FEB-2002.
XX
XX 01-AUG-2001; 2001WO-US024031.
XX
XX 02-AUG-2000; 2000US-0222599P.
XX 11-AUG-2000; 2000US-0224360P.
XX 11-APR-2001; 2001US-0282850P.
XX
XX (UYJO) UNIV JOHNS HOPKINS.
XX
XX St Croix B, Kinzler KW, Vogelstein B;
XX WPI; 2002-291856/33.
XX N-PSDB; ABL92104.
XX
PT An isolated molecule comprising an antibody variable region which
PT specifically binds to an extracellular domain of a tumor endothelial
PT marker (TEM) protein, useful for inhibiting tumor growth.
XX
XX Claim 1; Page 209-210; 33pp; English.
XX
XX The invention relates to an isolated molecule comprising an antibody
XX variable region which specifically binds to an extracellular domain of a
XX tumour endothelial marker (TEM) protein selected from ABB90732, ABB90740,
XX ABB90749, ABB90750 and ABB90769. The antibodies which bind to TEM
XX proteins have cytostatic, immunostimulant and antiangiogenic activity.
XX They are useful for inhibiting tumour growth, neovascularization in subjects
XX bearing a vascularised tumour, polycystic kidney disease, diabetic
XX retinopathy, rheumatoid arthritis and psoriasis. Human, mouse and rat TEM
XX genes and the encoded proteins (ABL92075-ABL92141 and ABB90721-ABB90789)
XX are disclosed, as are marker oligonucleotide sequences: tumour
XX endothelial markers (TEM) ABL91996-ABL92041 and ABL92143-ABL92191; normal
XX endothelial markers (NEM) ABL92042-ABL92074; and pan-endothelial markers
XX (PEM) ABL91903-ABL91995
XX
SQ Sequence 564 AA;
Query Match 100.0%; Score 1536; DB 5; Length 564;
Best Local Similarity 100.0%; Pred. No. 3e-156;
Matches 295; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGGRRDGGPACVGGFDLYFLDKSGSVLHNMNIYFYEQLAHKFISPOLRMSFIYFS 60
DB 27 GGGGRRDGGPACVGGFDLYFLDKSGSVLHNMNIYFYEQLAHKFISPOLRMSFIYFS 86
QY 61 TRGTTMLKLTEDREIQIGLEBELQKVLPGDVTYHBEGERASBOIYYENROGYRTASVII 120

Db 87 TRGTTLMKLTEDREIQRLGELQVLPGGDTVMHEGFERSSEQIYYENRQGYRTASVIT 146
Qy 121 ALTDGELHEDLFFYSEREARNSRDLGAIYVCVGYKDFNETQLARIADSKDHVPVNDGFO 180
Db 147 ALTDGELHEDLFFYSEREARNSRDLGAIYVCVGYKDFNETQLARIADSKDHVPVNDGFO 206
Qy 181 ALGGIHSILKSSCIEIIAABPSTTCAGESFOVVRNGGFRHARNDVRLCSFKINDSVT 240
Db 207 ALGGIHSILKSSCIEIIAABPSTTCAGESFOVVRNGGFRHARNDVRLCSFKINDSVT 266
Qy 241 LNEKPSVEDTYLLCPAPILKEVGKALQVSNMDGLSFISSVIITTHCSDGS 295
Db 267 LNEKPSVEDTYLLCPAPILKEVGKALQVSNMDGLSFISSVIITTHCSDGS 321

RESULT 6
ABR90724
ID ABR90724 standard; protein; 564 AA.
AC ABR90724;
XX
DT 30-MAY-2002 (first entry)
XX
DE Human Tumour Endothelial Marker polypeptide SEQ ID NO 187.
XX
KW Human; mouse; rat; TEM; tumour endothelial marker; NEM; PEM; cytostratic;
KW normal endothelial marker; pan-endothelial marker; immunostimulant;
KW antiangiogenic; tumour; neovascularisation; vascularised tumour;
KW polycystic kidney disease; diabetes; retinopathy; rheumatoid arthritis;
KW peoriaseis.
XX
OS Homo sapiens.
XX
PN W0200210217-A2.
XX
PD 07-FEB-2002.
XX
PF 01-AUG-2001; 2001WO-US024031.
XX
PR 02-AUG-2000; 2000US-0222599P.
XX 11-AUG-2000; 2000US-0224360P.
XX 11-APR-2001; 2001US-0282850P.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
XX
PI St Croix B, Kinzler KW, Vogelstein B;
XX
DR WPI; 2002-291856/33.
XX
XX An isolated molecule comprising an antibody variable region which
PT specifically binds to an extracellular domain of a tumor endothelial
PT marker (TEM) protein, useful for inhibiting tumor growth.
XX
PS Disclosure; Page 136-137; 331pp; English.
XX
XX The invention relates to an isolated molecule comprising an antibody
CC variable region which specifically binds to an extracellular domain of a
CC tumour endothelial marker (TEM) protein selected from ABR90732, ABR90740,
CC ABR90749, ABR90750 and ABR90769. The antibodies which bind to TEM
CC proteins have cytostratic, immunostimulant and antiangiogenic activity.
CC They are useful for inhibiting tumour growth, neovascularisation in subjects
CC bearing a vascularised tumour, polycystic kidney disease, diabetic
CC retinopathy, rheumatoid arthritis and psoriasis. Human, mouse and rat TEM
CC genes and the encoded proteins (ABR92075-ABR92141 and ABR90721-ABR90789)
CC are disclosed, as are marker oligonucleotide sequences: tumour
CC endothelial markers (TEM) ABR91996-ABR92041 and ABR92143-ABR92191; normal
CC endothelial markers (NEM) ABR92042-ABR92074; and pan-endothelial markers
CC (PEM) ABR91903-ABR91995
XX
XX Sequence 564 AA;
XX
Query Match 100.0%; Score 1536; DB 5; Length 564;

Best Local Similarity 100.0%; Pred. No. 3e-156;
Matches 295; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGGRRDDGPGACGDFLFIIDKSGSVLHHNNEIYYFVQLAHKFTSPQLRMSFVFS 60
Db 27 GGGRRDDGPGACGDFLFIIDKSGSVLHHNNEIYYFVQLAHKFTSPQLRMSFVFS 86
Qy 61 TRGTTLMKLTEDREIQRLGELQVLPGGDTVMHEGFERSSEQIYYENRQGYRTASVIT 120
Db 87 TRGTTLMKLTEDREIQRLGELQVLPGGDTVMHEGFERSSEQIYYENRQGYRTASVIT 146
Qy 121 ALTDGELHEDLFFYSEREARNSRDLGAIYVCVGYKDFNETQLARIADSKDHVPVNDGFO 180
Db 147 ALTDGELHEDLFFYSEREARNSRDLGAIYVCVGYKDFNETQLARIADSKDHVPVNDGFO 206
Qy 181 ALGGIHSILKSSCIEIIAABPSTTCAGESFOVVRNGGFRHARNDVRLCSFKINDSVT 240
Db 207 ALGGIHSILKSSCIEIIAABPSTTCAGESFOVVRNGGFRHARNDVRLCSFKINDSVT 266
Qy 241 LNEKPSVEDTYLLCPAPILKEVGKALQVSNMDGLSFISSVIITTHCSDGS 295
Db 267 LNEKPSVEDTYLLCPAPILKEVGKALQVSNMDGLSFISSVIITTHCSDGS 321

RESULT 7
ABP54904
ID ABP54904 standard; protein; 564 AA.
XX
AC ABP54904;
XX
DT 08-JAN-2003 (first entry)
XX
DE Human anthrax toxin receptor.
XX
KW Anthrax; toxin; receptor; human; TEM8; antibacterial.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 1..27
FT /label= Signal_peptide
FT 28..564
FT /label= Mature_protein
FT 28..320
FT /note= "extracellular domain"
FT 44..216
FT /note= "von Willebrand factor A domain"
FT 320..343
FT /note= "putative transmembrane domain"
FT 344..564
FT /note= "cytoplasmic domain"
XX
PN W0200246228-A2.
XX
PD 13-JUN-2002.
XX
PF 03-OCT-2001; 2001WO-US030941.
XX
PR 05-DEC-2000; 2000US-0251481P.
XX
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Young JAT, Bradley KA, Collier RJ, Mogridge JS;
XX
DR WPI; 2002-713235/77.
XX
XX N-PSDB; ABV73882.
XX
XX Novel isolated polypeptide useful for identifying agent that prevents or
PT reduces effect of anthrax toxin on host cell, for treating human or non-
PT human animal suffering from anthrax.
XX
XX Claim 1; Page 37-39; 45pp; English.
XX

CC The present sequence is the protein sequence of a human surface-bound
 CC anthrax toxin receptor (ATR), previously designated TEM8, and identified
 CC following a database screening using a newly isolated human surface-bound
 CC ATR (see ABP54903). The present sequence differs from the newly isolated
 CC human ATR only in the cytoplasmic domain (221 rather than 25 amino acids
 CC long), suggesting differential splicing of a primary mRNA transcript.
 CC Anthrax toxin protective antigen (PA) binds to these ATRs at a von
 CC Willebrand factor A domain located in the extracellular domain of ATR.
 CC The invention provides ATR polypeptides and polynucleotides, vectors,
 CC host cells, and transgenic and knock-out animals. It also provides the
 CC methods for identifying molecules that bind the ATR and which reduce the
 CC toxicity of anthrax toxin. A claimed method for treating anthrax in a
 CC human or animal involves administering an agent that inhibits binding
 CC between PA and ATR at a level effective to reduce the severity of
 CC anthrax. Suitable agents include the present polypeptide or a PA-binding
 CC fragment of it, a PA-binding polypeptide at least 80% identical to these,
 CC a fusion protein, a monoclonal or polyclonal antibody, a polysaccharide,
 CC a lipid or a nucleic acid

XX Sequence 564 AA:

Query Match 100.0%; Score 1536; DB 5; Length 564;
 Best Local Similarity 100.0%; Pred. No. 3e-156;
 Matches 295; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGRRREDGGPACGCGFDLYFILDKSGSVLHNNHNIYFVYEQLAHKFISPOLKMSFIVFS 60
 DB 27 GGGRRREDGGPACGCGFDLYFILDKSGSVLHNNHNIYFVYEQLAHKFISPOLKMSFIVFS 86
 QY 61 TRGTTLMKLTEDREIQROGLEBLQKVLPGDITVHGEFERASQIYYENRGYRTASVIT 120
 DB 87 TRGTTLMKLTEDREIQROGLEBLQKVLPGDITVHGEFERASQIYYENRGYRTASVIT 146
 QY 121 ALTDGELHEDLFFYSERANRSRDIGAIVYCVGVDPFNETQLARIADSKDHVFPVNDGFO 180
 DB 147 ALTDGELHEDLFFYSERANRSRDIGAIVYCVGVDPFNETQLARIADSKDHVFPVNDGFO 206
 QY 181 ALGGIHSILKSCIEILAEPTTICAGESFOVVVRNGGFRHARNDRVLCSPKINDSVT 240
 DB 207 ALGGIHSILKSCIEILAEPTTICAGESFOVVVRNGGFRHARNDRVLCSPKINDSVT 266
 QY 241 LNEKPSVEDTYLLCPAPILKEVGKALQVSMNDGLSFSSSVIITTHCSDGS 295
 DB 267 LNEKPSVEDTYLLCPAPILKEVGKALQVSMNDGLSFSSSVIITTHCSDGS 321

RESULT 8
 ABUS4457
 ID ABUS4457 standard; protein; 564 AA.

XX AC ABUS4457;
 XX DT 12-MAR-2003 (first entry)

XX DE Human tumour endothelial marker TEM 19.

XX KW Human; endothelial cell; EC; tumour endothelial cell; TEM; NEM;
 KW Tumour endothelial marker; normal endothelial marker; PEM;
 KW pan-endothelial marker; polycystic kidney disease; psoriasis;
 KW diabetic retinopathy; rheumatoid arthritis; tumour angiogenesis;
 KW neovascularization; immune response; cytostatic; antidiabetic;
 KW ophthalmological; antirheumatic; antiarthritic; antipsoriatic.

XX OS Homo sapiens.

XX PN WO200283874-A2.

XX PD 24-OCT-2002.

XX PF 10-APR-2002; 2002WO-US008253.

XX PR 11-APR-2001; 2001US-0282850P.

XX PR 06-FEB-2002; 2002US-0354262P.

XX (UYJO) UNIV JOHNS HOPKINS.
 PA Carson-Walter E, St Croix B, Kinzler KW, Vogelstein B;
 XX WPI, 2003-093016/08.
 XX N-PSDB; ABX72029.
 DR N-PSDB; ABX72029.
 XX New purified human transmembrane protein, designated as tumor endothelial
 PT marker (TEM) 3, useful for detecting, diagnosing or treating tumors,
 PT polycystic kidney disease, diabetic retinopathy, rheumatoid arthritis or
 PT psoriasis.

XX Disclosure; Page 226-227; 374pp; English.

XX The present invention relates to a novel method for the isolation of
 CC endothelial cells (ECs), and the identification of genes expressed in
 CC normal and tumour ECs. Tumour endothelial marker (TEM) normal
 CC endothelial marker (NEM), and pan-endothelial marker (PEM) genes are
 CC identified in human ECs. The human EC marker proteins and the
 CC polynucleotide sequences encoding them are useful for detecting,
 CC diagnosing or treating tumours as well as polycystic kidney disease,
 CC diabetic retinopathy, rheumatoid arthritis, and psoriasis. They are also
 CC useful for inhibiting neovascularization or tumour angiogenesis, for
 CC inducing an immune response to tumour endothelial cells in a patient, or
 CC for identifying candidate drugs for treating tumours. The present
 CC sequence represents a human TEM or NEM protein of the invention

XX Sequence 564 AA:

Query Match 100.0%; Score 1536; DB 6; Length 564;
 Best Local Similarity 100.0%; Pred. No. 3e-156;
 Matches 295; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGRRREDGGPACGCGFDLYFILDKSGSVLHNNHNIYFVYEQLAHKFISPOLKMSFIVFS 60
 DB 27 GGGRRREDGGPACGCGFDLYFILDKSGSVLHNNHNIYFVYEQLAHKFISPOLKMSFIVFS 86
 QY 61 TRGTTLMKLTEDREIQROGLEBLQKVLPGDITVHGEFERASQIYYENRGYRTASVIT 120
 DB 87 TRGTTLMKLTEDREIQROGLEBLQKVLPGDITVHGEFERASQIYYENRGYRTASVIT 146
 QY 121 ALTDGELHEDLFFYSERANRSRDIGAIVYCVGVDPFNETQLARIADSKDHVFPVNDGFO 180
 DB 147 ALTDGELHEDLFFYSERANRSRDIGAIVYCVGVDPFNETQLARIADSKDHVFPVNDGFO 206
 QY 181 ALGGIHSILKSCIEILAEPTTICAGESFOVVVRNGGFRHARNDRVLCSPKINDSVT 240
 DB 207 ALGGIHSILKSCIEILAEPTTICAGESFOVVVRNGGFRHARNDRVLCSPKINDSVT 266
 QY 241 LNEKPSVEDTYLLCPAPILKEVGKALQVSMNDGLSFSSSVIITTHCSDGS 295
 DB 267 LNEKPSVEDTYLLCPAPILKEVGKALQVSMNDGLSFSSSVIITTHCSDGS 321

RESULT 9
 ABUS4431
 ID ABUS4431 standard; protein; 564 AA.

XX AC ABUS4431;
 XX DT 12-MAR-2003 (first entry)

XX DE Human tumour endothelial marker TEM 8.

XX KW Human; endothelial cell; EC; tumour endothelial cell; TEM; NEM;
 KW Tumour endothelial marker; normal endothelial marker; PEM;
 KW pan-endothelial marker; polycystic kidney disease; psoriasis;
 KW diabetic retinopathy; rheumatoid arthritis; tumour angiogenesis;
 KW neovascularization; immune response; cytostatic; antidiabetic;
 KW ophthalmological; antirheumatic; antiarthritic; antipsoriatic.

XX OS Homo sapiens.

XX XX WO200283874-A2.
XX XX 24-OCT-2002.
XX XX 10-APR-2002; 2002WO-US008253.
XX XX 11-APR-2001; 2001US-0282850P.
XX XX 06-FEB-2002; 2002US-0354262P.
XX XX (UNIV) UNIV JOHNS HOPKINS.
XX XX Carson-Walter E, St Croix B, Kinzler KW, Vogelstein B;
XX XX WPI, 2003-093016/08.
XX XX
XX XX New purified human transmembrane protein, designated as tumor endothelial
XX XX marker (TEM) 3, useful for detecting, diagnosing or treating tumors,
XX XX polycystic kidney disease, diabetic retinopathy, rheumatoid arthritis or
XX XX psoriasis.
XX XX
XX XX Disclosure; Page 138-139; 374pp; English.
XX XX
XX XX The present invention relates to a novel method for the isolation of
XX XX endothelial cells (ECs), and the identification of genes expressed in
XX XX normal and tumor ECs. Tumour endothelial marker (TEM), normal
XX XX endothelial marker (NEM), and pan-endothelial marker (PEM) genes are
XX XX identified in human ECs. The human EC marker proteins and the
XX XX polynucleotide sequences encoding them are useful for detecting,
XX XX diagnosing or treating tumors as well as polycystic kidney disease,
XX XX diabetic retinopathy, rheumatoid arthritis, and psoriasis. They are also
XX XX useful for inhibiting neoangiogenesis or tumour angiogenesis, for
XX XX inducing an immune response to tumour endothelial cells in a patient, or
XX XX for identifying candidate drugs for treating tumours. The present
XX XX sequence represents a human TEM or NEM protein of the invention
XX XX
XX XX Sequence 564 AA;
XX XX
XX XX Query Match 100.0%; Score 1536; DB 6; Length 564;
XX XX Best Local Similarity 100.0%; Pred. No. 3e-156;
XX XX Matches 295; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGRRREDGGPACYGFDLYFTLDKSGSVLHNMNIYYFVEQLAKFTSPQLRMSFVFS 60
DB 27 GGGRRREDGGPACYGFDLYFTLDKSGSVLHNMNIYYFVEQLAKFTSPQLRMSFVFS 86
QY 61 TRGTTLMKLTDRRQIQGLELOKVLPGGDTYMHGFERASQIYYENRGYRTASYII 120
DB 87 TRGTTLMKLTDRRQIQGLELOKVLPGGDTYMHGFERASQIYYENRGYRTASYII 146
QY 121 ALTDGELHEDLFFYSEREARNSRDGAIYVCVGVDFNETQLARIADSKDHFVPVNDGFO 180
DB 147 ALTDGELHEDLFFYSEREARNSRDGAIYVCVGVDFNETQLARIADSKDHFVPVNDGFO 206
QY 181 ALGGIHSILKKSCTEILAAEPSTICAGESPQVVVRNGGFRHARNDVRLCSFKINDSVT 240
DB 207 ALGGIHSILKKSCTEILAAEPSTICAGESPQVVVRNGGFRHARNDVRLCSFKINDSVT 266
QY 241 LNEKFSVEDTYLLCPAPILKEVKAKALQVSMNDGLSFISSVITTTTSCSDGS 295
DB 267 LNEKFSVEDTYLLCPAPILKEVKAKALQVSMNDGLSFISSVITTTTSCSDGS 321
XX XX
XX XX RESULT 10
XX XX ADJ70017
XX XX ADJ70017 standard; protein; 564 AA.
XX XX
XX XX ADJ70017;
XX XX
XX XX 06-MAY-2004 (first entry)
XX XX
XX XX Human heat mitochondrial protein as a therapeutic target SeqID1823.
XX XX

KW mitochondrial; human; screening assay; diabetes mellitus;
KW Huntington's disease; osteoarthritis;
KW Leber's hereditary optic neuropathy; LHON;
KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
KW neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;
KW osteopathic; ophthalmological; cyostatic.
XX XX
XX XX Homo sapiens.
XX XX
XX XX WO2003087768-A2.
XX XX
XX XX 23-OCT-2003.
XX XX
XX XX 04-APR-2003; 2003WO-US010870.
XX XX
XX XX 12-APR-2002; 2002US-0372843P.
XX XX 17-JUN-2002; 2002US-0389987P.
XX XX 20-SEP-2002; 2002US-0412418P.
XX XX
XX XX (MITO-) MITOKOR.
XX XX (BUCK-) BUCK INST AGE RES.
XX XX
XX XX Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM,
XX XX Warnock DE;
XX XX WPI, 2003-845369/78.
XX XX
XX XX Identifying a mitochondrial target for drug screening assays and for
XX XX treating diseases associated with altered mitochondrial function,
XX XX comprises detecting a modified polypeptide in a sample and correlating
XX XX with the disease.
XX XX
XX XX Claim 1; SEQ ID NO 1823; 180pp; English.
XX XX
XX XX This invention relates to novel mitochondrial targets that can be used
XX XX for therapeutic intervention in treating a disease associated with
XX XX altered mitochondrial function. Specifically, it refers to a method for
XX XX identifying proteins of the human heart mitochondrial proteome that are
XX XX useful for drug screening assays, as well as therapeutic targets. The
XX XX present invention describes a method for identifying such proteins that
XX XX can be used in the treatment of various diseases associated with altered
XX XX mitochondrial function including diabetes mellitus, Huntington's disease,
XX XX osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
XX XX encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
XX XX ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
XX XX compositions have neuroprotective, nootropic, antidiabetic,
XX XX anticonvulsant, antiarthritic, osteopathic, ophthalmological and
XX XX cyostatic activities. This polypeptide sequence is a human heart
XX XX mitochondrial protein of the invention.
XX XX
XX XX Sequence 564 AA;
XX XX
XX XX Query Match 100.0%; Score 1536; DB 7; Length 564;
XX XX Best Local Similarity 100.0%; Pred. No. 3e-156;
XX XX Matches 295; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGRRREDGGPACYGFDLYFTLDKSGSVLHNMNIYYFVEQLAKFTSPQLRMSFVFS 60
DB 27 GGGRRREDGGPACYGFDLYFTLDKSGSVLHNMNIYYFVEQLAKFTSPQLRMSFVFS 86
QY 61 TRGTTLMKLTDRRQIQGLELOKVLPGGDTYMHGFERASQIYYENRGYRTASYII 120
DB 87 TRGTTLMKLTDRRQIQGLELOKVLPGGDTYMHGFERASQIYYENRGYRTASYII 146
QY 121 ALTDGELHEDLFFYSEREARNSRDGAIYVCVGVDFNETQLARIADSKDHFVPVNDGFO 180
DB 147 ALTDGELHEDLFFYSEREARNSRDGAIYVCVGVDFNETQLARIADSKDHFVPVNDGFO 206
QY 181 ALGGIHSILKKSCTEILAAEPSTICAGESPQVVVRNGGFRHARNDVRLCSFKINDSVT 240
DB 207 ALGGIHSILKKSCTEILAAEPSTICAGESPQVVVRNGGFRHARNDVRLCSFKINDSVT 266
XX XX

Qy 241 LNEKPSVEDTYLLCPAPILKEVGKALQVSMNDGLSFSSSVITTTTCSDS 295
Db 267 LNEKPSVEDTYLLCPAPILKEVGKALQVSMNDGLSFSSSVITTTTCSDS 321

RESULT 11
ID ADR48216 standard; protein; 564 AA.
AC ADR48216;
XX 18-NOV-2004 (first entry)
DT
DE Human tumour endothelial marker 8 precursor protein SEQ:4.
XX pancreatic cancer-associated transcript; pancreatic cancer; human;
KW cytostatic; gene therapy; protein therapy;
KW tumour endothelial marker 8 precursor; TEM8.
OS Homo sapiens.
XX MO2004074510-A1.
XX 02-SEP-2004.
XX 18-FEB-2004; 2004WO-AU000194.
XX 18-FEB-2003; 2003AU-00900747.
XX (GARV-) GARVAN INST MEDICAL RES.
XX Biankin A, Segara D, Henshall S, Sutherland R;
PI WPI; 2004-635591/61.
DR N-PSDB; ADR48215.

PT Detecting pancreatic cancer-associated transcript in a biological sample,
PT useful for diagnosing or treating the disease, comprises contacting the
PT sample with a polynucleotide that selectively hybridizes to a specific
PT sequence.

XX Claim 70; SEQ ID NO 4; 263pp; English.

CC The present invention describes a method for detecting a pancreatic
CC cancer-associated transcript in a biological sample. The method comprises
CC contacting the biological sample with a polynucleotide that selectively
CC hybridizes to a sequence at least 80% identical to a sequence as shown in
CC any one of Tables 3 to 25 in the specification or having the GenBank
CC Accession Number AF279145. Also described: (1) diagnosing pancreatic
CC cancer in a human or animal subject being tested, determining the
CC likelihood that a subject having a pancreatic cancer will survive, or
CC determining the suitability of a subject having a pancreatic cancer for
CC surgical resection therapy; (2) detecting a pancreatic cancer-associated
CC polypeptide in a biological sample; (3) determining the likelihood that a
CC subject having a pancreatic cancer will survive; and (4) monitoring the
CC efficacy of a therapeutic treatment of pancreatic cancer. A pancreatic
CC cancer-associated transcript has cytostatic activity, and can be used in
CC gene and protein therapy. A pancreatic cancer-associated transcript
CC polynucleotide, a vector comprising the polynucleotide, an isolated
CC polypeptide or an antibody that binds to the isolated polypeptide can be
CC used for diagnosing or prognosing pancreatic cancer or for preparing a
CC medicament for the treatment of pancreatic cancer. The prognostic or
CC diagnostic methods are useful for the early detection of pancreatic
CC cancer or its metastases, and for monitoring the progress of disease such
CC as during remission or following surgery or chemotherapy. The present
CC sequence represents human tumour endothelial marker 8 precursor (TEM8),
CC which is used in the exemplification of the present invention.

XX Sequence 564 AA;
SQ

Query Match 100.0%; Score 1536; DB 8; Length 564;
Best Local Similarity 100.0%; Pred. No. 3e-156;
Matches 295; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGRRRDGGPACGCGFDLYFIIDKSGSVLHNNNIYYFVEQLAHKEISPOLRMSFIYFS 60
Db 27 GGGRRRDGGPACGCGFDLYFIIDKSGSVLHNNNIYYFVEQLAHKEISPOLRMSFIYFS 86

Qy 61 TRGTTLMKLTEDRROIRQGLEELQKVLPGDDTYMHGEFERASQIYYENRQYRTASVII 120
Db 87 TRGTTLMKLTEDRROIRQGLEELQKVLPGDDTYMHGEFERASQIYYENRQYRTASVII 146

Qy 121 ALTDGEIHEDLFFYSREARSRDLGAIYCVGVKDFNETOLARIADSKDHFVPNDGFO 180
Db 147 ALTDGEIHEDLFFYSREARSRDLGAIYCVGVKDFNETOLARIADSKDHFVPNDGFO 206

Qy 181 ALQGIHSILKSCIEILAAEPSTTCAGSFOVVVRNGFRHARNDVILCSFKINDSVT 240
Db 207 ALQGIHSILKSCIEILAAEPSTTCAGSFOVVVRNGFRHARNDVILCSFKINDSVT 266

Qy 241 LNEKPSVEDTYLLCPAPILKEVGKALQVSMNDGLSFSSSVITTTTCSDS 295
Db 267 LNEKPSVEDTYLLCPAPILKEVGKALQVSMNDGLSFSSSVITTTTCSDS 321

RESULT 12
AD100550 standard; protein; 551 AA.
ID AD100550;
AC AD100550;
XX 22-APR-2004 (first entry)
DT
DE Human TANGO 197 Ig fusion mutated protein - plasmid p0610.
XX fusion; von Willebrand factor A-like domain; vWF; antibacterial;
KW cutaneous; inhalation anthrax; human; TANGO 197 Ig fusion; mutant;
KW plasmid p0610; mutein.

OS Homo sapiens.
XX Synthetic.
XX US2003144193-A1.
XX 31-JUL-2003.
XX 24-JUL-2002; 2002US-00201292.
XX 20-DEC-2001; 2001US-00038307.
XX (ROTT/) ROTTMAN J B.
PA (OKEE/) O'KEEFE T L.
PA (OZKA/) OZKAYNAK E.
PA (HEAL/) HEALEY J J.
PI Rottman JB, O'Keefe TL, Ozkaynak E, Healey JJ;
XX WPI; 2003-720708/68.
DR N-PSDB; AD100549.

PT New TANGO 197 and/or TANGO 216 fusion polypeptides for preventing or
PT ameliorating symptoms of anthrax comprises a von Willebrand factor A-like
PT domain (vWF) amino acid sequence and an amino acid sequence heterologous
PT to the vWF.

XX Claim 45; SEQ ID NO 18; 86pp; English.

CC The invention relates to a novel fusion polypeptide comprising a von
CC Willebrand factor A-like domain (vWF) amino acid sequence and an amino
CC acid sequence heterologous to the vWF. The polypeptide of the invention
CC demonstrates antibacterial activities whilst the composition and method
CC may be useful in preventing or ameliorating the symptoms of cutaneous
CC and/or inhalation anthrax. The current sequence is that of the human
CC TANGO 197 Ig mutated fusion protein of the invention.

XX Sequence 551 AA;
SQ

Query Match 99.7%; Score 1532; DB 7; Length 551;
 Best Local Similarity 100.0%; Pred. No. 7.8e-156;
 Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGRRREDGGPACGCGFDLYFLIDKSGSVLHHNMEIYFVBOIAHKFTSPQLRMSFIYFS 60
 DB 27 GGGRRREDGGPACGCGFDLYFLIDKSGSVLHHNMEIYFVBOIAHKFTSPQLRMSFIYFS 86
 QY 61 TRGTTLMKLTEDREQIRQGLEBELQKVLPGDPTVMHEGFERASEQIYYENRGYRTASVII 120
 DB 87 TRGTTLMKLTEDREQIRQGLEBELQKVLPGDPTVMHEGFERASEQIYYENRGYRTASVII 146
 QY 121 ALTDGELHEDLFFYSEREARNSRDIGAIYVCVGVDFNETQLARIADSKDHVPVNDGFQ 180
 DB 147 ALTDGELHEDLFFYSEREARNSRDIGAIYVCVGVDFNETQLARIADSKDHVPVNDGFQ 206
 QY 181 ALQGIHSHILKKSCEILAAEPSTTCAGESFOVVVRNGRFRHARNDVRLCSFKINDSVT 240
 DB 207 ALQGIHSHILKKSCEILAAEPSTTCAGESFOVVVRNGRFRHARNDVRLCSFKINDSVT 266
 QY 241 LNEKPFSEVDYTLCPAPILKEVGKKAALQVSNMDGLSFISSVYITTHCSDG 294
 DB 267 LNEKPFSEVDYTLCPAPILKEVGKKAALQVSNMDGLSFISSVYITTHCSDG 320

RESULT 13
 ADM64576
 ID ADM64576 standard; protein; 551 AA.
 AC ADM64576;
 XX 03-JUN-2004 (first entry)
 DT
 XX Mouse TANGO197-immunoglobulin (Ig) fusion protein.
 DE
 XX antibacterial; gene therapy;
 KM von Willebrand factor A-like domain amino acid sequence;
 KM vWF amino acid sequence; anthrax; Bacillus anthracis; cutaneous anthrax;
 KM inhalation anthrax; mouse; TANGO197; immunoglobulin; Ig; fusion protein;
 KM gene.
 XX Mus sp.
 OS Synthetic.
 OS
 XX US2003134786-A1.
 PN
 XX 17-JUL-2003.
 PD
 XX 20-DEC-2001; 2001US-00038307.
 PF
 XX 20-DEC-2001; 2001US-00038307.
 PR
 XX 20-DEC-2001; 2001US-00038307.
 XX
 PA (ROTT/) ROTTMAN J B.
 PA (OKEE/) O'KEEFE T L.
 PA (OZKA/) OZKAYNAK E.
 PA (HEAL/) HEALEY J J.
 XX
 PI Rottman JB, O'keefe TL, Ozkaynak E, Healey JJ;
 XX
 XX WPI, 2003-829643/77.
 DR N-PSDB; ADM64575.
 DR
 XX
 PT New TANGO 197 and/or TANGO 216 fusion polypeptides for preventing or
 PT ameliorating symptoms of anthrax comprises a von Willebrand factor A-like
 PT domain (vWF) amino acid sequence and an amino acid sequence heterologous
 PT to the vWF.
 XX
 XX Claim 44; SEQ ID NO 18; 64pp; English.
 PS
 CC The invention describes a fusion polypeptide comprising a von Willebrand
 CC factor A-like domain (vWF) amino acid sequence and an amino acid sequence
 CC heterologous to the vWF. Also described are: a method of preventing or

CC ameliorating a symptom of anthrax in a subject thought to be at risk for
 CC exposure to or suspected of having been exposed to Bacillus anthracis;
 CC and a pharmaceutical composition comprising the novel fusion polypeptide.
 CC The composition and method are useful in preventing or ameliorating
 CC symptoms of cutaneous and/or inhalation anthrax. This is the amino acid
 CC sequence of a fusion protein comprising mouse TANGO197 and immunoglobulin
 CC (Ig) that can be used to treat exposure to or prevent a symptom of
 CC anthrax.
 CC
 XX Sequence 551 AA:
 XX

Query Match 99.7%; Score 1532; DB 7; Length 551;
 Best Local Similarity 100.0%; Pred. No. 7.8e-156;
 Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGRRREDGGPACGCGFDLYFLIDKSGSVLHHNMEIYFVBOIAHKFTSPQLRMSFIYFS 60
 DB 27 GGGRRREDGGPACGCGFDLYFLIDKSGSVLHHNMEIYFVBOIAHKFTSPQLRMSFIYFS 86
 QY 61 TRGTTLMKLTEDREQIRQGLEBELQKVLPGDPTVMHEGFERASEQIYYENRGYRTASVII 120
 DB 87 TRGTTLMKLTEDREQIRQGLEBELQKVLPGDPTVMHEGFERASEQIYYENRGYRTASVII 146
 QY 121 ALTDGELHEDLFFYSEREARNSRDIGAIYVCVGVDFNETQLARIADSKDHVPVNDGFQ 180
 DB 147 ALTDGELHEDLFFYSEREARNSRDIGAIYVCVGVDFNETQLARIADSKDHVPVNDGFQ 206
 QY 181 ALQGIHSHILKKSCEILAAEPSTTCAGESFOVVVRNGRFRHARNDVRLCSFKINDSVT 240
 DB 207 ALQGIHSHILKKSCEILAAEPSTTCAGESFOVVVRNGRFRHARNDVRLCSFKINDSVT 266
 QY 241 LNEKPFSEVDYTLCPAPILKEVGKKAALQVSNMDGLSFISSVYITTHCSDG 294
 DB 267 LNEKPFSEVDYTLCPAPILKEVGKKAALQVSNMDGLSFISSVYITTHCSDG 320

RESULT 14
 AAE01469
 ID AAE01469 standard; protein; 403 AA.
 AC AAE01469;
 XX 17-JUL-2001 (first entry)
 DT
 XX Human gene 4 encoded secreted protein HMLFR02, SEQ ID NO:125.
 DE
 XX Human; secreted protein; proliferative disorder; cancer; tumour;
 KM foetal abnormality; developmental abnormality; haematopoietic disorder;
 KM immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KM inflammation; allergy; neurological disorder; Alzheimer's disease;
 KM Parkinson's disease; cognitive disorder; schizophrenia; asthma;
 KM skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
 KM cardiovascular disorder; angiogenic disorder; kidney disorder;
 KM gastrointestinal disorder; pregnancy-related disorder;
 KM endocrine disorder; infection; wound healing; vulnerability; cell culture;
 KM chemotaxis; food additive; gene therapy; binding partner identification;
 KM chromosome 19.
 XX
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT 1..27
 FT Peptide /label= signal_peptide
 FT Protein 28..403
 FT /note= "Mature human secreted protein"
 FT Misc-difference 175
 FT /label= Unknown
 FT /note= "Encoded by GKT"
 FT Misc-difference 320
 FT /label= Unknown
 FT /note= "Encoded by SGT"
 FT Misc-difference 331
 FT /label= Unknown

FT	Misc-difference	/note= "Encoded by KTC"
FT	368	
FT	/label= Unknown	
XX	/note= "Encoded by WGC"	
PX	MOZ00134626-AI.	
PN		
PD	17-MAY-2001.	
PP		
PR	01-NOV-2000; 2000OMO-US030045.	
PR	05-NOV-1999; 99US-0163581P.	
PR	30-JUN-2000; 2000OUS-0215133P.	
PA	(HUMA-) HUMAN GENOME SCI INC.	
PI	Ruben SM, Komatsoulis GA, Moore PA, Birse CE, Ni J;	
DR	WPI; 2001-308778/32.	
DR	N-P5DB; AAD05334.	
PT	New nucleic acid molecules encoding 28 human secreted proteins for	
PT	diagnosing, preventing, treating or ameliorating medical conditions and	
PT	used as food additives or preservatives.	
PS	Claim 11; Page 505-506; 562pp; English.	
CC	AAD05300-AAD05379 represent cDNAs corresponding to 28 human secreted	
CC	protein genes, and AAE01436-AAE01513 represent the proteins they encode.	
CC	AAE01514-AAE01544 represent human secreted protein fragments or variants.	
CC	The genes and their secreted proteins are useful for preventing, treating,	
CC	or ameliorating medical conditions, e.g., by protein or gene therapy.	
CC	Pathological conditions can be diagnosed by determining the amount of the	
CC	new protein in a sample or by determining the presence of mutations in	
CC	the new genes. Specific uses are described for each of the 28 genes,	
CC	based on the tissues in which they are most highly expressed, and include	
CC	developing products for the diagnosis or treatment of proliferative	
CC	disorders, cancer, tumours, foetal and developmental abnormalities,	
CC	haematopoietic disorders, diseases of the immune system, AIDS, autoimmune	
CC	diseases (e.g., rheumatoid arthritis), inflammation, allergies,	
CC	neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),	
CC	cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,	
CC	psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,	
CC	angiogenic disorders, kidney disorders, gastrointestinal disorders,	
CC	pregnancy-related disorders, endocrine disorders, and infections. The	
CC	proteins can also be used to aid wound healing and epithelial cell	
CC	proliferation, to prevent skin aging due to sunburn, to maintain organs	
CC	before transplantation, for supporting cell culture of primary tissues,	
CC	to regenerate tissues, to identify their cognate ligands or binding	
CC	partners, and in chemotherapy, and can be used as a food additive or	
CC	preservative to modify storage properties. Antibodies specific for a	
CC	protein of the invention can be used in alleviating symptoms associated	
CC	with the disorders mentioned above, and in diagnostic immunoassays e.g.,	
CC	radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The	
CC	present sequence represents a human secreted protein of the invention	
XX		
SQ	Sequence 403 AA:	
Query Match	99.2%; Score 1524; DB 4; Length 403;	
Best Local Similarity	99.3%; Pval. No. 3.5e-155;	
Matches 293; Conservative	0; Mismatches 2; Indels 0; Gaps 0	
DB	1 GGGGRBDGPGPCYGGFDLYFLIDSGSVLHHMNEIYYVEVDLAHKFISPOLRMPSTVS 60	
27	GGGGRBDDGPACIGCFDLYFLIDSGSVLHMNEIYYVEVDLAHKFISPOLRMPSTVS 86	
QY	TRGTTLMKLTEDREOIRGOILEIKOVLPFGDITYMEHGFERASBOIYYENROGRTASVII 120	
DB	TRGTTLMKLTEDREOIRGOILEIKOVLPFGDITYMEHGFERASBOIYYENROGRTASVII 146	
QY	ALTDEGLHNDLEFFYSREANRSKDGAIIYCVGVDFNFTOLARIKSDKHVPVNVDSGQ 180	
DB	ALTDELHNDLEFFYSREANRSKDGAIIYCVGVDFNFTOLARIKSDKHVPVNVDSGQ 206	

QY	181	ALAGIHSLIKSKSCTEITLAAEPSTICGSESFQVVRNNGRPHANVRYLCSFPIINDSVT	240
DB	207	ALAGIHSLIKSKSCTEITLAAEPSTICGSESFQVVRNNGRPHANVRYLCSFPIINDSVT	266
QY	241	INERKPSVEDTYLLCPAPILKEVGKAAQVSNMNDGLSTISSVLIITTHCSDGS	295
DB	267	INERKPSVEDTYLLCPAPILKEVGKAAQVSNMNDGLSTISSVLIITTHCSDGS	321
RESULT 15			
ID	ABG63873	standard; protein; 403 AA.	
XX	ABG63873;		
AC			
DT	27-AUG-2002	(first entry)	
XX			
DE		Human albumin fusion protein #548.	
XX			
KM		Albumin fusion protein; therapeutic protein X; human albumin; HA;	
KM		human serum albumin; HSA; cancer; reproductive disorder;	
KM		digestive disorder; immune disorder; endocrine disorder;	
KM		haematopoietic disorder; neural disorder; connective disorder;	
KM		cystostatic; antifertility; antiinflammatory; antitumor;	
KM		immunomodulator; anti-HIV; antidiabetic; haemostatic; neutropic;	
KM		neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;	
KM		osteopathic; antiarthritic.	
OS		Homo sapiens.	
OS		Synthetic.	
XX			
XX	WO200177137-A1.		
PD			
XX	18-OCT-2001.		
PD			
XX	12-APR-2001; 2001WO-US011988.		
PF			
XX	12-APR-2000; 2000US-0229358P.		
PR	25-APR-2000; 2000US-0199384P.		
PR	21-DEC-2000; 2000US-0256931P.		
XX			
PA	(HUMA-) HUMAN GENOME SCI INC.		
PI			
XX	Rosen CA, Haseltine WA;		
XX	WPI: 2002-010886/01.		
DR			
XX			
PT		New fusion protein for treating disease e.g. diabetes comprises an	
XX		albumin fused to a therapeutic protein.	
PS			
XX	Claim 1; Page 872-873; 2102pp; English.		
CC		The present invention relates to albumin fusion proteins comprising a	
CC		therapeutic protein X and human albumin (HA, also known as human serum	
CC		albumin, HSA). The proteins are useful for treating a disease or disorder	
CC		that may be modulated by therapeutic protein X. The albumin extends the	
CC		shelf-life of protein X, and may increase its biological in vitro/in vivo	
CC		activity. The protein is useful for treating and diagnosing disorders	
CC		such as cancer, reproductive disorders, digestive disorders (e.g. Crohn's	
CC		disease, ulcerative colitis), immune disorders (e.g. acquired	
CC		immunodeficiency syndrome, AIDS), endocrine disorders (e.g. diabetes),	
CC		haematopoietic disorders, neural disorders (e.g. Alzheimer's,	
CC		Parkinson's, Creutzfeldt-Jacob disease, encephalomyelitis, meningitis,	
CC		schizophrenia), and connective disorders (e.g. osteoporosis, arthritis).	
CC		ABG63326-ABG65518 represent albumin fusion proteins of the invention	
XX			
XX	Sequence 403 AA;		
Query Match	99.2%	Score 1524; DB 5; Length 403;	
Best Local Similarity	99.3%	Pred. No. 3.5e-155;	
Matches 293; Conservative	0; Mismatches 2;	Indels 0; Gaps 0;	

```

Qy 1 GGGRRRDEGSPACYGFDLYFIIDKSGSVLHMHNEIYFVEQLAHKFI SPQLRMSFIYFS 60
   |||||
Db 27 GGGRRRDEGSPACYGFDLYFIIDKSGSVLHMHNEIYFVEQLAHKFI SPQLRMSFIYFS 86
   |||||
Qy 61 TRGTTLMKLTEDREQIRQGLELELOKVLPGSDTYMHGEPERASEQIYYENRQGYRTASVTI 120
   |||||
Db 87 TRGTTLMKLTEDREQIRQGLELELOKVLPGSDTYMHGEPERASEQIYYENRQGYRTASVTI 146
   |||||
Qy 121 ALTDGELHEDLFPYSEREANRSRDIGAIYVCVGVDFNETQLARIADSKDHVPVNDGFQ 180
   |||||
Db 147 ALTDGELHEDLFPYSEREANRSRDIGAIYVCVGVDFNETQLARIADSKDHVPVNDGFQ 206
   |||||
Qy 181 ALQGIHSHILKKSCEIILAEPSSTICAGESFOVVVRGNGFRHANNVDRVLCSPKINDSVT 240
   |||||
Db 207 ALQGIHSHILKKSCEIILAEPSSTICAGESFOVVVRGNGFRHANNVDRVLCSPKINDSVT 266
   |||||
Qy 241 LNEKPFVSVEDTYLCPAPILKEVGKALQVSMNDGLSFISSVIITTHCSDGS 295
   |||||
Db 267 LNEKPFVSVEDTYLCPAPILKEVGKALQVSMNDGLSFISSVIITTHCSDXS 321
   |||||

```

Search completed: June 13, 2005, 19:56:46
 Job time : 111.221 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 19:37:47 ; Search time 99.0301 Seconds

(without alignments)
1525,429 Million cell updates/sec

Title: US-09-970-076-2_COPY_27_321

Perfect score: 1536

Sequence: 1 GGGRRRDGPGACGFGFDLY.....GLSFSSSVITTTTTCSDGS 295

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1536	100.0	564	1 ATRI_HUMAN	Q96X22 homo sapien
2	1519	98.9	562	1 ATRI_MOUSE	Q96Z52 mus musculu
3	790.5	51.5	489	1 ATR2_HUMAN	P58335 homo sapien
4	770	50.1	487	2 G6DFX2	G6DFX2 mus musculu
5	606	39.5	641	2 O8BVM2	O8BVM2 mus musculu
6	388	25.3	97	2 Q96EC6	Q96EC6 homo sapien
7	159.5	10.4	1161	1 ITAD_RAT	Q96Y67 rattus norv
8	156	10.2	1162	1 ITAD_HUMAN	Q13349 homo sapien
9	153.5	10.0	1332	2 Q9BPO8	Q9BPO8 halocynthia
10	148.5	9.7	1626	2 O8NFM1	O8NFM1 mus musculu
11	148	9.6	441	2 O8T5C3	O8T5C3 mytilus edu
12	148	9.6	444	2 O8T5C3	O8T5C3 mytilus edu
13	145.5	9.5	1152	1 ITAM_HUMAN	P11215 homo sapien
14	145	9.4	453	2 O8T5C2	O8T5C2 mytilus gal
15	143.5	9.3	1163	1 ITAX_HUMAN	P20702 homo sapien
16	143	9.3	1124	1 CA1C_CHICK	P13944 gallus gall
17	141	9.2	1153	1 ITAM_MOUSE	P05555 mus musculu
18	139	9.0	724	2 Q04588	Q04588 elmeria max
19	139	9.0	3119	1 CA1C_MOUSE	Q06844 mus musculu
20	134.5	8.8	712	2 O43981	O43981 elmeria ten
21	134	8.7	760	2 O70350	O70350 mus musculu
22	132.5	8.6	765	2 O9U8J9	O9U8J9 neospora ca
23	131.5	8.6	517	2 O43853	O43853 homo sapien
24	131	8.5	760	1 CO2_MOUSE	P21180 mus musculu
25	131	8.5	3063	1 CA1C_HUMAN	Q95715 homo sapien
26	130.5	8.5	920	2 Q28984	Q28984 sus scrofa
27	130	8.4	1888	1 CA1E_CHICK	P32018 gallus gall
28	129.5	8.4	1182	2 O8C6F9	O8C6F9 mus musculu
29	128	8.3	221	2 O99L17	O99L17 mus musculu
30	128	8.3	3183	2 Q65ZC2	Q65ZC2 caenorhabdi
31	128	8.3	3767	1 MUA3_CAEEL	P34576 caenorhabdi

32	123	8.0	1151	2 Q9J130	Q9J130 rattus norv
33	121	7.9	758	2 Q6M373	Q6M373 rattus norv
34	121	7.9	758	2 Q8CIP8	Q8CIP8 rattus norv
35	121	7.9	2401	2 Q7RF52	Q7RF52 plasmodium
36	120.5	7.8	790	2 Q6DC06	Q6DC06 xenopus lae
37	118.5	7.7	440	2 O8C8T2	O8C8T2 mus musculu
38	118.5	7.7	848	2 O8C720	O8C720 mus musculu
39	118.5	7.7	919	2 Q7SR52	Q7SR52 lymanaea sta
40	118.5	7.7	3567	2 Q9ES77	Q9ES77 mus musculu
41	118	7.7	652	2 Q95L12	Q95L12 bos taurus
42	118	7.7	929	1 CA1C_NOTVI	Q91145 notophthalm
43	117.5	7.6	637	2 O8IVY1	O8IVY1 homo sapien
44	117.5	7.6	1284	2 Q6P159	Q6P159 homo sapien
45	117.5	7.6	1329	1 KF10_HUMAN	Q9P218 homo sapien

ALIGNMENTS

RESULT 1

ID	ATRI_HUMAN	STANDARD;	PRT;	564 AA.
AC	Q96X22:096P02:Q9NVP3:			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	05-JUN-2004 (Rel. 44, Last annotation update)			
DE	Anthrax toxin receptor 1 precursor (Tumor endothelial marker 8).			
GN	Name=ANTXR1; Synonyms=ATR, TEM8;			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RX	MEDLINE=20407466; PubMed=10947988; DOI=10.1126/science.289.5482.1197;			
RA	St Croix B., Rago C., Velculescu V., Traverso G., Romans K.B.,			
RA	Montgomery B., Lai A., Riggs G.J., Lengauer C., Vogelstein B.,			
RA	Kinzler K.W.;			
RT	"Genes expressed in human tumor endothelium";			
RL	Science 289:1197-1202(2000)..			
RN	[2]			
RP	SEQUENCE FROM N.A. (ISOFORM 2), AND INTERACTION WITH ANTHRAX TOXIN.			
RX	MEDLINE=21557240; PubMed=11700562; DOI=10.1038/n5101998;			
RA	Bradley K.A., Mogridge J., Mourez M., Collier R.J., Young J.A.T.;			
RT	"Identification of the cellular receptor for anthrax toxin.";			
RL	Nature 414:225-229(2001).			
RN	[3]			
RP	SEQUENCE OF 184-564 FROM N.A. (ISOFORM 1), AND SEQUENCE FROM N.A.			
RP	(ISOFORM 3).			
RX	PubMed=14702039; DOI=10.1038/ng1285;			
RA	Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,			
RA	Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,			
RA	Sekine M., Odayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,			
RA	Yanamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,			
RA	Nagahara K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,			
RA	Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,			
RA	Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,			
RA	Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,			
RA	Yamazaki M., Nishimura K., Ishibashi T., Yamashita H., Murakawa K.,			
RA	Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoaka S., Chiba Y.,			
RA	Ishida S., Ono Y., Takiguchi S., Watanabe S., Yoshida M., Hotta T.,			
RA	Kuano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,			
RA	Nomura Y., Togiyasu S., Komai F., Hara R., Takeuchi K., Arita M.,			
RA	Iinose N., Mutsaers K., Yuki H., Ohshima A., Sasaki N., Aotsuka S.,			
RA	Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,			
RA	Moriya S., Momiyama H., Satoh N., Takami S., Terashima F., Suzuki O.,			
RA	Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,			
RA	Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,			
RA	Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,			
RA	Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujizawa T.,			
RA	Ono T., Yamada K., Fujii Y., Ozaki K., Hiro M., Omori Y.,			
RA	Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,			
RA	Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senda T.,			

RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Matanabe Y., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuno Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.,
RA "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RT Nat. Genet. 36:40-45 (2004).
RN [4]
RN SEQUENCE FROM N.A. (ISOFORM 4).
RC TISSUE=Kidney;
RC MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RX Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Ditschenko L., Marulanda K., Farmer A.A., Rubin G.M., Hong L.,
RA Stepheon M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carrinci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywinski M.I., Skalski U., Smalhus D.E.,
RA Schercher A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [5]
RN INTERACTION WITH ANTHRAX TOXIN.
RC TISSUE=Placenta;
RX MEDLINE=22608610; PubMed=12700348; DOI=10.1073/pnas.0431098100;
RA Scoble H.M., Rainey G.U.A., Bradley K.A., Young J.R.T.,
RT "Human capillary morphogenesis protein 2 functions as an anthrax toxin
RT receptor.";
RT Proc. Natl. Acad. Sci. U.S.A. 100:5170-5174 (2003).
RN [6]
RN SPLICING ISOFORM (S) THAT ARE POTENTIAL NMD TARGET (S).
RX PubMed=14759238; DOI=10.1186/gb-2004-5-2-r8;
RA Hillman R.T., Green R.B., Brenner S.E.,
RT "An unappreciated role for RNA surveillance.";
RT Genome Biol. 5:RESEARCH008.1-RESEARCH008.16 (2004).
CC -1- FUNCTION: Cellular role is not yet known.
CC -1- SUBUNIT: Binds to the protective antigen (PA) of *Bacillus*
CC anthracis. Binding does not occur in the presence of calcium.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Probable).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Comment=Experimental confirmation may be lacking for some
CC isoforms;
CC Name=1;
CC IsoId=Q9H6X2-1; Sequence=Displayed;
CC Note=May be produced at very low levels due to a premature stop
CC codon in the mRNA, leading to nonsense-mediated mRNA decay;
CC Name=2;
CC IsoId=Q9H6X2-2; Sequence=VSP_000444, VSP_000445;
CC Name=3;
CC IsoId=Q9H6X2-3; Sequence=VSP_000446, VSP_000447;
CC Name=4;
CC IsoId=Q9H6X2-4; Sequence=VSP_000448, VSP_000449;
CC -1- TISSUE SPECIFICITY: Highly expressed in tumor endothelial cells
CC but not in normal endothelial cells
CC -1- DOMAIN: Binding to PA seems to be effected through the VWA domain.
CC -1- SIMILARITY: Belongs to the ATR family.
CC -1- SIMILARITY: Contains 1 VWA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF279145; AAK52094.1; -
DR EMBL: AF421380; AAL26496.1; -
DR EMBL: AK025429; BAB15128.1; ALT_INIT.
DR EMBL: AK001463; BAA91707.1; ALT_FRAME.
DR EMBL: BC012074; AAI12074.1; -
DR Genew: HGNC:21014; ANTXRL.
DR H-InvDB: HIX0002125; -
DR MIM: 606410; -
DR InterPro: IPR008400; Anth_Ig.
DR InterPro: IPR008399; Ant_C.
DR InterPro: IPR020315; VWF_A.
DR Pfam: PF05587; Anth_Ig; 1.
DR Pfam: PF00092; VWA; 1.
DR Pfam: PF05586; Ant_C; 1.
DR SMART: SM00327; VWA; 1.
DR PROSITE: PS50234; VWF; 1.
KW Alternative splicing; Glycoprotein; Receptor; Signal; Transmembrane.
FT SIGNAL 1 32
FT CHAIN 3 564
FT DOMAIN 33 321
FT TRANSMEM 322 342
FT DOMAIN 343 564
FT DOMAIN 44 215
FT DOMAIN 360 368
FT DOMAIN 506 564
FT CARBOHYD 166 166
FT CARBOHYD 184 184
FT CARBOHYD 262 262
FT VARSPPLIC 365 368
FT VARSPPLIC 369 564
FT VARSPPLIC 268 297
FT VARSPPLIC 298 564
FT VARSPPLIC 319 333
FT VARSPPLIC 334 564
FT SEQUENCE 564 AA; 62789 MW; B18A00AD5DF2233 CRC64;
Query Match 100.0%; Score 1536; DB 1; Length 564;
Best Local Similarity 100.0%; Pred. No. 2,4e-114;
Matches 295; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGRRRDEGGPACGPGDLYFTLIDKSGSVLAHNNETYYFPEOLAKRFPOLRMSFVFS 60
DB 27 GGGRRRDEGGPACGPGDLYFTLIDKSGSVLAHNNETYYFPEOLAKRFPOLRMSFVFS 86
QY TRGTTLTKLLEDEBQIQGLEIQLVPGSDTYMHGEPASBQIYYENRQGYRTASVII 120
DB 87 TRGTTLTKLLEDEBQIQGLEIQLVPGSDTYMHGEPASBQIYYENRQGYRTASVII 146
QY 121 ALTDGEIHEDLFFYSREANRSDLGAIYVCVGKDFNETQLARIADSKDHVPVNDGFO 180
DB 147 ALTDGEIHEDLFFYSREANRSDLGAIYVCVGKDFNETQLARIADSKDHVPVNDGFO 206
QY 181 ALGGIHSILKSGCETILAEPTICGSEFOVVVNGRGRHARNDRVLCSPFINDSVT 240
DB 207 ALGGIHSILKSGCETILAEPTICGSEFOVVVNGRGRHARNDRVLCSPFINDSVT 266
QY 241 INKRFSEVDYTLCPAPILKEVGMKALQVSNMNGLSFSSSVIITTHCSDS 295
DB 267 INKRFSEVDYTLCPAPILKEVGMKALQVSNMNGLSFSSSVIITTHCSDS 321

ATRI_MOUSE	STANDARD	PRT	562 AA.
AC	09C252		
DT	28-FEB-2003 (Rel. 41, Created)		
DT	28-FEB-2003 (Rel. 41, Last sequence update)		
DT	05-JUL-2004 (Rel. 44, Last annotation update)		
DE	Anchorin toxin receptor 1 precursor (Tumor endothelial marker 8).		
GN	Name=Anchorin; Synonym=Alt, Tem6;		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A. (ISOFORM 1).		
RX	MEBLIN=21443268; PubMed=11559328;		
RA	Carson-Walter E.B., Wacklin D.N., Nanda A., Vogelestein B.,		
RT	Kinzler K.W., St Croix B.;		
RT	"Cell surface tumor endothelial markers are conserved in mice and		
RL	humans."		
RL	Cancer Res. 61:6649-6655(2001).		
RN	[2]		
RP	SEQUENCE OF 88-562 FROM N.A. (ISOFORM 2).		
RP	STRAIN=C57BL/6J; TISSUE=Embryo;		
RX	MEBLIN=22354683; PubMed=12466651; DOI=10.1038/nature01266;		
RA	Okezaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,		
RA	Mikaido I., Oeato N., Saito K., Suzuki H., Yamahata I., Kiyosawa H.,		
RA	Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,		
RA	Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,		
RA	Schmitt L.M., Knapin A., Matsuda H., Batalov S., Beisel K.W.,		
RA	Blake J.A., Bradt D., Brusic V., Chochia C., Corbani L.E., Cousins S.,		
RA	Dalla E., Dragan T.A., Fletcher C.P., Forrest A., Frerz K.S.,		
RA	Gaesteland T., Gariboldi M., Giesi C., Godzik A., Gough J.,		
RA	Grimmond S., Guestindich S., Hirokawa N., Jackson I.J., Jarvis E.D.,		
RA	Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., Kling B.L.,		
RA	Konagaya A., Kurochkin I.V., Lee Y., Lemhard B., Lyons P.A.,		
RA	Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,		
RA	Mageshima T., Numata K., Okido T., Pavan W.J., Petrea G., Peoole G.,		
RA	Petrovsky N., Pillai R., Pontius U.U., Qi D., Ramchandran S.,		
RA	Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,		
RA	Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,		
RA	Suterana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,		
RA	Verardo R., Wagner L., Wahlestedt C., Wang Y., Matanabe Y., Wells C.,		
RA	Wilmig L.G., Wyrshaw-Boris A., Yangisawa M., Yang I., Yang L.,		
RA	Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carrinci P., Hayatsu N.,		
RA	Hirozane-Fukushima T., Kono H., Nakamura M., Sakazume N., Sato K.,		
RA	Shiraki T., Waki K., Kawai J., Atzawa K., Arakawa T., Fukuda S.,		
RA	Hara A., Hashizume W., Imocant K., Jihli Y., Itoh M., Kagawa I.,		
RA	Myazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,		
RA	Yasunishi A., Yoshino M., Waterson R., Lander E.S., Rogers J.,		
RA	Briney E., Hayashizaki Y.;		
RT	"Analysis of the mouse transcriptome based on functional annotation of		
RT	60,770 full-length cDNAs."		
RL	Nature 420:563-573(2002).		
CC	-1- FUNCTION: Cellular role is not yet known.		
CC	-1- SUBUNIT: Binds to the protective antigen (PA) of Bacillus anthracis (By similarity).		
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein (Probable).		
CC	-1- ALTERNATIVE PRODUCTS:		
CC	Event=Alternative splicing; Named isoforms=2;		
CC	Name=1;		
CC	IsoId=09C252-1; Sequence=Displayed;		
CC	Name=2;		
CC	IsoId=09C252-2; Sequence=VSD_000450;		
CC	Note=No experimental confirmation available;		
CC	-1- DOMAIN: Binding to PA seems to be effected through the VWA domain (By similarity).		
CC	-1- SIMILARITY: Belongs to the ATR family.		
CC	-1- SIMILARITY: Contains 1 VWA domain.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	-----		

CC		the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed.
CC		modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.ebi.ac.uk/announcement/) or send an email to license@ebi-eib.ch .
CC		-----
DR	EMBL; AF378762; A111999.1; -	-
DR	EMBL; AK013005; BAB28591.1; ALT_INT1.	
DR	MGI; MGI:1916788; Antxr1.	
DR	InterPro; IPR008400; Anth_Ig.	
DR	InterPro; IPR008399; Ant_C.	
DR	InterPro; IPR002035; VWF_A.	
DR	Pfam; PF05587; Anth_Ig; I.	
DR	Pfam; PF05586; Ant_C; I.	
DR	Pfam; PF00092; VWA; 1.	
DR	PROSITE; PS50234; VWFA; 1.	
KW	Alternative splicing; Glycoprotein; Receptor; Signal; Transmembrane.	
FT	SIGNAL	1-30 potential.
FT	CHAIN	31-562 Antitoxin receptor 1.
FT	DOMAIN	31-319 Extracellular (Potential).
FT	TRANSMEM	320-340 Potential.
FT	DOMAIN	341-562 Cytoplasmic (Potential).
FT	DOMAIN	42-213 VWF.
FT	DOMAIN	358-365 Asp/Glu-rich (highly acidic).
FT	CARBOHYD	501-562 Pro-rich.
FT	CARBOHYD	164-164 N-linked (GLNAC. . .) (Potential).
FT	CARBOHYD	182-182 N-linked (G1CNAC. . .) (Potential).
FT	CARBOHYD	260-260 N-linked (GLCNAc. . .) (Potential).
FT	VANSPIC	477-562 GRCHFRVNSQAPKPLNNTHPSPPAPPTPPAP HCPPPASAPTPPIPSPPSTLPPPPQAPPPRAPPSPRP RPST -> RFRGMVLITCLDSKVHPRHDKGPETLUKA WMFSSFLERAFO (in Isoform 2). /FTID=VSP_000450.
SO	SEQUENCE	562 AA; 62308 MW; 6AC92045B4BB4FC CRC64;
Query Match	Best Local Similarity	98.9%; Score 1519; DB 1; Length 562;
Matches 291; Conservative		Pred. No. 5.5e-113; 2; Mismatches 2; Indels 0; Gaps 0;
Dy	1	GCGGRREDGGPACGCGFDLYFLIDKSGSVLHHNNEIYFPYEOLAHKEISPOLRMSPFVFS 60
Dy	25	GHGRRRDGPACTGCAGDFLYFLIDKSGSVLHHNNEIYFPYEOLAHKEISPOLRMSPFVFS 84
Dy	61	TGTGTLTKLTEDREQLRQGLELOKVLPGSDTYMHGEFPASEQIYYENRGYRTASVII 120
Dy	85	TGTGTLTKLTEDREQLRQGLELOKVLPGSDTYMHGEFPASEQIYYENRGYRTASVII 144
Dy	121	ALTDGEIHEDLPFYSEERANSRDGLAIYVCVGKDENEQTOLARIADSKOHVPVNDGFQ 180
Dy	145	ALTDEGEIHEDLPFYSEERANSRDGLAIYVCVGKDENEQTOLARIADSKOHVPVNDGFQ 204
Dy	181	ALOGIIHSILKSCIEITLAEPEPTICGESFOVVVRNGRPHARNDRVLCSPFIINDSYT 240
Dy	205	ALOGIIHSILKSCIEITLAEPEPTICGESFOVVVRNGRPHARNDRVLCSPFIINDSYT 264
Dy	241	LNKPFEVEDTYLLCPAPILKEVGMKALOVSMNDGLSFISSSYYITTTCSDDGS 295
Dy	265	LNKPFRAVEDTYLLCPAPILKEVGMKALOVSMNDGLSFISSSYYITTTCSDDGS 319
RESULT 3		
ATR2_HUMAN	ID	NAME
AC	P58335; Q86U11; Q8NB13; Q96NC7;	STANDARD; PRT; 489 AA.
DT	28-FEB-2003 (Rel. 41, Created)	
DT	10-OCT-2003 (Rel. 42, Last sequence update)	
DT	05-JUL-2004 (Rel. 44, Last annotation update)	
DE	Antitoxin toxin receptor 2 precursor (Capillary morphogenesis protein-2) (CMG-2).	
GN	Name=ANTXR2; Synonyms=CMG2;	
OS	Homo sapiens (human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	

OX NCBI_TaxId=9606;
 [1]
 RN SEQUENCE FROM N.A. (ISOFORM 2).
 RP MEDLINE=21539596; PubMed=11693410;
 RX Bell S.E., Mavila A., Salazar R., Bayless K.J., Kanagala S.,
 RA Maxwell S.A., Davis G.E.;
 RT "Differential gene expression during capillary morphogenesis in 3D
 RT collagen matrices: regulated expression of genes involved in basement
 RT membrane matrix assembly, cell cycle progression, cellular
 RT differentiation and G-protein signaling.";
 RL J. Cell Sci. 114:2755-2773(2001).
 [2]
 RN SEQUENCE FROM N.A. (ISOFORM 1), AND INTERACTION WITH ANTHRAX TOXIN.
 RP TISSUE=Placenta;
 RX MEDLINE=22606610; PubMed=12700348; DOI=10.1073/pnas.0431098100;
 RA Scobie H.M., Rainey G.J.A., Bradley J.A., Young J.A.T.;
 RT "Human capillary morphogenesis protein 2 functions as an anthrax toxin
 RT receptor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:5170-5174(2003).
 [3]
 RN SEQUENCE OF 78-489 FROM N.A. (ISOFORM 3), AND SEQUENCE FROM N.A.
 RP (ISOFORM 4).
 RX TISSUE=Synovial cell;
 PubMed=14702039; DOI=10.1038/ng1285;
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Wakematsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sekine M., Ohyashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
 RA Nagahori K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
 RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
 RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
 RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
 RA Yamazaki M., Nishimura K., Ishibashi T., Yamashita H., Murakawa K.,
 RA Fujimori K., Tanai H., Kimata S., Matsumabe M., Hirooka S., Chiba Y.,
 RA Ishida S., Ono Y., Takiguchi S., Matsumabe S., Yoshida M., Hoota T.,
 RA Kusano Y., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
 RA Nomura Y., Togiyasu S., Komai F., Hara R., Takeuchi K., Arita M.,
 RA Imobe N., Mutsaers K., Yuki H., Oshima A., Sasaki N., Aotsuka S.,
 RA Yoshihawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mitoiguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hibiakaki S., Matsumabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujitani T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiroo M., Omori T.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Oktani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Togoashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita K.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs.";
 RL Nat. Genet. 36:40-45(2004).
 CC -1- FUNCTION: Cellular role is not yet known.
 CC -1- SUBUNIT: Binds to the protective antigen (PA) of Bacillus
 CC anthracis in a divalent cation-dependent manner, with the
 CC following preference: calcium > manganese > magnesium > zinc.
 CC Seems to bind to collagen type IV and laminin.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (isoforms 1 and 2).
 CC Secreted (isoform 3). Isoform 1 is expressed at the cell surface
 CC while isoform 2 is predominantly expressed within the endoplasmic
 CC reticulum and not at the plasma membrane.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=4;
 CC Name=1;
 CC IsoId=P58335-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=P58335-2; Sequence=VSP_008343;
 CC Name=3;
 CC IsoId=P58335-3; Sequence=VSP_008344, VSP_008345;
 CC Note=No experimental confirmation available;

```

CC      Name=4;
CC      Ircid=P58335-4; Sequence=VSP_008346;
CC      Note-Non experimental confirmation available;
CC      -1- TISSUE SPECIFICITY: Expressed in colon, heart, kidney, lung,
CC      liver, peripheral blood leukocytes, placenta, skeletal muscle,
CC      small intestine and spleen.
CC      -1- DOMAIN: Binding to PA seems to be effected through the VWA domain.
CC      -1- SIMILARITY: Belongs to the ATR family.
CC      -1- SIMILARITY: Contains 1 VWFA domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC      or send an email to license@ebi.ac.uk).
CC      -----
DR      EMBL; AY040326; AAK7222.1; -
DR      EMBL; AY233452; AAP04016.1; -
DR      EMBL; AK056536; BAB70976.1; ALT_ININ.
DR      EMBL; AK091721; BAC03731.1; -
DR      Genew; HGNC:21732; ANTXR2.
DR      MIM; 608041; -
DR      InterPro; IPR008399; Ant_C.
DR      InterPro; IPR020355; VWF_A.
DR      Pfam; PF05586; Ant_C.1.
DR      Pfam; PF00092; VWA; 1.
DR      PROSITE; PS50234; VWFA; 1.
KW      Alternative splicing; Glycoprotein; Receptor; Signal; Transmembrane.
FT      SIGNAL 1 33
FT      CHAIN 34 489
FT      DOMAIN 34 318
FT      TRANSMEM 319 341
FT      DOMAIN 342 489
FT      DOMAIN 44 213
FT      CARBOHYD 250 250
FT      CARBOHYD 260 260
FT      VASPLIC 213 315
FT      VASPLIC 290 322
FT      VASPLIC 323 489
FT      VASPLIC 477 489
FT      VASPLIC 489 AA; 53692 MM; B9F679DB75BE2B7 CRC64;
FT      CONFLICT 357 357
FT      SEQUENCE 489 AA; 53692 MM; B9F679DB75BE2B7 CRC64;
SQ
Query Match 51.5%; Score 790.5; DB 1; Length 489;
Best Local Similarity 53.2%; Pred. No. 8.9e-55;
Matches 157; Conservative 54; Mismatches 81; Indels 3; Gaps 2;
QY      1 GGG-REDDGPAVYGGFDLYEILIDKSGSVLHHNNELIYFVEQLAHKRFISPLMSFVIF 59
DB      26 GGGELRPAQEPQSCRPAFDLIFYVDKSGSVANNHIEIYNFVQDLERFVSPENKLSFIVF 85
QY      60 STGTITMLKLTDEDEQIIRQGLIELQKLYLPGGDTMHGSGFERASBOIYYENRQYRTASVI 119
DB      86 SSGQITILPLTGDGRGKISKGLIEDLKRVSPVGETYIHHGLKLANQI--QKAGGLKTSII 143
QY      120 IALTDGSLHEDLFFYSREANRSGDLGAIYVCVGAKPNEGQLARIADSKDHVPVANDGF 179
DB      144 IALTDGKLDGLVPSYAEKKAETISLSIASYCVGVLPFEOQLERIDASKQVFPVKGGF 203
QY      180 QALGGIHSILKSKCIEILAAEPSTICAGSFFQVAVVNGNGRRHARANDRVLSKFSKINDSV 239
DB      204 QALGGIINSILAGSCTEILQLSPSSVCGVEFQVLVLSGRGMLGSRNSVLCITYAVET 263
QY      240 TLNRKPSVEDDYLLCPAPILKEVGMKALQVSNMDGLSFISSSVILITTHCSDG 294

```


Db 264 TTSVAKPSVQJNSMLCPAPILNKAGETLDSVSPFGKGSVTSGLITVATCSNG 318

RESULT 4

06DFX2 PRELIMINARY; PRT; 487 AA.

AC Q6DFX2; 25-OCT-2004 (TREMBlrel. 28, Created)

DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)

DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)

DE Anthrax toxin receptor 2.

GN Name=Anthr2;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6; TISSUE=Eye;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marziani K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshlycki S., Carninci P., Prange C., Raha S.S., Loguclano N.A., Peters G.J., Abramson R.D., Mulhaly S.J., Boak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richardson S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Pehey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grilwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinska M.I., Skalek U., Smallegange D.E., Scherch A., Schein J.B., Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";

RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6; TISSUE=Eye;

RA Strauberg R.;

RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.

RX EMBL: BC076595; AAF76595.1;

DR GO: 0016021; C: integral to membrane; IEA.

DR GO: 0004872; F: receptor activity; IEA.

DR InterPro: IPR008400; Ant_19.

DR InterPro: IPR008399; Ant_G.

DR InterPro: IPR02035; VWF_A.

DR Pfam: PF05587; Anth_1g1.

DR Pfam: PF05586; Ant_C_1.

DR Pfam: PF00092; VMA_1.

DR SMART: SM00327; VMA_1.

DR PROSITE: PS50234; VMA_1.

DR Receptor.

SO SEQUENCE 487 AA; 53184 MW; 61A400D60BCD69 CRC64;

Query Match 50.1%; Score 770; DB 2; Length 487;

Best Local Similarity 52.8%; Pred. No. 3.9e-53;

Matches 150; Conservative 52; Mismatches 80; Indels 2; Gaps 1;

Qy 11 PACGGGDIYPIIDKSSVLMHMEITYYPVQGLAKHTISPOLRMSFVPSRGITLMKLT 70

Db 37 PSCKKADLVYFDKSSGVANWMIETYNFVQLTERFVSPMRISFTVSSQATIIILPT 96

Qy 71 EDREQIQGLEELQKVPFGDPTVWHEGFERASEQIYVENRGYRTASVITALTGGELED 130

Db 97 GDRKIKIGKLEDLKAVPFGVETTYHKGKLANEQI--QNAQGLKASSIIITLDTGKLDGL 154

Qy 131 LFPYSREANRSDLGAIYVCVGVKDFNETQLARIADSKDHVPVNDGFOALQGIHSIL 190

Db 155 VPSYAENBAKKSRLGASVYCVGVLDFEQALERIASKQOVFPVKGFOALKGIINSIL 214

Qy 191 KKSCTEILAAEPSTICGSEFSFOVVNRGNFRHANVDRVCSFKNDSVTLNKRPFSVD 250

Db 215 AQSTTEILELSPSSVCGEFPQVLTGRATYSISHDSVLCFTFANSTYTKSEKPVSIQ 274

Qy 251 TYLLCPAPILKEVGMKAALQVSMNDGLSFISSSVYITTTTCSDG 294

Db 275 SSILCPAPVINKQGETLEVGISVYDGSVSRSLTITATCTGNG 318

RESULT 5

08BVM2 PRELIMINARY; PRT; 641 AA.

AC 08BVM2;

DT 01-MAR-2003 (TREMBlrel. 23, Created)

DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:493430J11 product:hypothetical Prolin-rich region/von Willebrand factor type A domain containing protein, full insert sequence.

GN Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Testis;

RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;

RA Carninci P., Hayashizaki Y.;

RT "High-efficiency full-length cDNA cloning.";

RT Meth. Enzymol. 303:19-44(1999).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Testis;

RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;

RA RIKEN FANTOM Consortium;

RT "Functional annotation of a full-length mouse cDNA collection.";

RT Nature 409:685-690(2001).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Testis;

RA The FANTOM Consortium.

RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";

RT Nature 420:563-573(2002).

RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Testis;

RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;

RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;

RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.";

RT Genome Res. 10:1617-1630(2000).

RN [5]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Testis;

RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;

RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P., Kono H., Akiyama U., Nishi K., Kiteunai T., Tashiro H., Itoh M., Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A., Yamamoto R., Matsunoto H., Sakaguchi S., Ikegami T., Kashiwagi K., Fujisake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsumura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;

RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer.";

RT Genome Res. 10:1757-1771(2000).

RN [6]

RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RA Adachi J., Aizawa K., Akimura T., Arikawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hasehita W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hirao T., Hirose T.,
 RA Hoti F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai Y., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akehira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
 RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK077206; BAC36683.1; -
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR InterPro; IPR006400; Arch_1g.
 DR InterPro; IPR02035; VWF_A.
 DR Pfam; PF05587; Ant1_1g; 1.
 DR Pfam; PF00092; VWA; 1.
 DR PRINTS; PRO0453; VMPADOMAIN.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS50234; VWA; 1.
 DR Hypothetical protein.
 KW SEQUENCE 641 AA; 70415 MW; 199E300730BC85E3 CRC64;

Query Match 39.5%; Score 606; DB 2; Length 641;
 Best Local Similarity 42.2%; Pred. No. 7,2e-40;
 Matches 119; Conservative 58; Mismatches 105; Indels 0; Gaps 0;

QY 10 GPACGDFLYFLIDKSGSVLHNNELIYFVGLAKHFKISFQLFMSFVSTGTTMLKL 69
 DB 68 GDDCGGIFDLVLDKSGSVADNWIHYSFAGELVKKFTNNELISITTSFAEVLPL 127
 QY 70 TEDREQIROGLEELQKVLPGDDTYMHGFEFASQIYENNGQRTASVITLADGEHLE 129
 DB 128 TSDSEKINKSLVLSKSVIPQGLTHMQGLRANQIRKSTLGGRIVNSVITLADGILL 187
 QY 130 DLPFYSEREARNSRDLGAIYVCVGVKDFNETOLARLADS KDHPFVNDGFOALGIIHSI 189
 DB 188 KPYLDITMEAKKARMGAIYTVTVGVFMYSSKQGLVNIAGDPDRCGVDEGFGALGVDP 247
 QY 190 LKKSCEITLAAPESTTCAGSFQVAVVRNGFRHARNDRVLCSPKINDSVTLNKPSPSVE 249
 DB 248 TSKECTEILSVQPTVYCAKDFYQVNIISGHGLNNTSNMKQVLCRFKFSKVDSPSDMN 307
 QY 250 DTVLLCPAPILKEVGMKALGVSMNDGLSFSSSVITTTTTC 291
 DB 308 EHSITCPGPKIKHTGEDVSLQVLSLNGISFTGNKLIITSTNC 349

RESULT 6

Q96EC6 PRELIMINARY; PRT; 97 AA.
 AC Q96EC6; 01-DEC-2001 (TREMBlrel. 19; Created)
 DT 01-DEC-2001 (TREMBlrel. 19; Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25; Last annotation update)
 DE ANTXR1 protein (Fragment).
 GN Name=ANTXR1;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Breast;
 RX MEDLINE=23388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold B.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner J., Schenker C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueda T.B., Toehiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abremson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez M.I., Skalko U., Smallos D.E., Schnerch A., Schein J.E.,
 RA Krzyzanski M.I., Skalko U., Smallos D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marz M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Breast;
 RA Strausberg R.;
 RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC012475; AAH12475.1; -
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR InterPro; IPR006400; Ant1_1g; 1.
 DR Pfam; PF05587; Ant1_1g; 1.
 DR NON TER 1
 PT 1
 SQ SEQUENCE 97 AA; 10453 MW; 14F475F0B170E71A CRC64;

Query Match 25.3%; Score 388; DB 2; Length 97;
 Best Local Similarity 95.1%; Pred. No. 1.9e-23;
 Matches 78; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 211 FQVVAENGFRHARNDRVLCSPKINDSVTLNKPSPSVEVDYLLCPAPILKEVGMKALQ 270
 DB 1 FQVVAENGFRHARNDRVLCSPKINDSVTLNKPSPSVEVDYLLCPAPILKEVGMKALQ 60
 QY 271 VSMNDGLSFSSSVITTTTTC 292
 DB 61 VSMNDGLSFSSSVITTTTTC 82

RESULT 7

ITAD_RAT STANDARD; PRT; 1161 AA.
 AC O90YER;
 DT 05-JUL-2004 (Rel. 44; Created)
 DT 05-JUL-2004 (Rel. 44; Last sequence update)
 DT 05-JUL-2004 (Rel. 44; Last annotation update)
 DE Integrin alpha-D precursor.
 GN Name=Itgad;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RA O'Brien M.M., VanderVlieten M., Klugmann P.D., Dietrich G.,
 RA Gallatin W.M.;
 RT "Cloning of rat alpha D, a novel beta 2 integrin."
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Integrin alpha-D/beta-2 is a receptor for ICAM3 and
 CC VCAM1. May play a role in the atherosclerotic process such as
 CC clearing lipoproteins from plaques and in phagocytosis of blood-
 CC borne pathogens, particulate matter, and senescent erythrocytes
 CC from the blood (By similarity).
 CC -1- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-D
 CC associates with beta-2 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
 CC -1- DOMAIN: The integrin I-domain (inset) is a VWA domain. Integrins
 CC with I-domains do not undergo protease cleavage.
 CC -1- SIMILARITY: Belongs to the integrin alpha chain family.

CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
CC -1- SIMILARITY: Contains 1 WFPA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF021334; AAF21241.1; -.
DR HSBP; P1215; IBO.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; WFPA.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_alpha; 1.
DR Pfam; PF00092; WFPA; 1.
DR PRINTS; PR00185; INTEGRINA.
DR PRINTS; PR00453; WFPA DOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VMA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; WFPA; 1.
DR Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
KW Repeat; Signal; Transmembrane.
FT SIGNAL 1 19 Potential.
FT CHAIN 20 1161 Integrin alpha-D.
FT DOMAIN 20 1100 Extracellular (Potential).
FT TRANSMEM 1101 1121 Potential.
FT DOMAIN 1122 1161 Cytoplasmic (Potential).
FT REPEAT 34 87 FG-GAP 1.
FT REPEAT 88 7 FG-GAP 2.
FT DOMAIN 152 334 WFPA.
FT REPEAT 352 402 FG-GAP 3.
FT REPEAT 403 454 FG-GAP 4.
FT REPEAT 456 517 FG-GAP 5.
FT REPEAT 519 577 FG-GAP 6.
FT REPEAT 582 634 FG-GAP 7.
FT CA_BIND 467 475 Potential.
FT CA_BIND 531 539 Potential.
FT CA_BIND 594 602 Potential.
FT SITE 1126 1130 GPFKR motif.
FT DISULFID 69 76 By similarity.
FT DISULFID 108 126 By similarity.
FT DISULFID 656 711 By similarity.
FT DISULFID 769 775 By similarity.
FT DISULFID 845 860 By similarity.
FT DISULFID 993 1017 By similarity.
FT DISULFID 1022 1027 By similarity.
FT CARBOHYD 61 61 N-linked (GLCNAc...) (Potential).
FT CARBOHYD 89 89 N-linked (GLCNAc...) (Potential).
FT CARBOHYD 245 245 N-linked (GLCNAc...) (Potential).
FT CARBOHYD 393 393 N-linked (GLCNAc...) (Potential).
FT CARBOHYD 696 696 N-linked (GLCNAc...) (Potential).
FT CARBOHYD 734 734 N-linked (GLCNAc...) (Potential).
FT CARBOHYD 784 784 N-linked (GLCNAc...) (Potential).
FT CARBOHYD 907 907 N-linked (GLCNAc...) (Potential).
FT CARBOHYD 936 936 N-linked (GLCNAc...) (Potential).
FT CARBOHYD 1045 1045 N-linked (GLCNAc...) (Potential).
SQ SEQUENCE 1161 AA; 126600 MW; 2258491A98A4705E CRC64;
Query Match 10.4%; Score 159.5; DB 1; Length 1161;
Best Local Similarity 27.9%; Pred. No. 0.00079;
Matches 63; Conservative 41; Mismatches 93; Indels 29; Gaps 11;
Qy 18 DLYPLDKSGSV-LHHNWEIYFVQOLAKFTISPOLMSFIVFSTRGTTLMKTE----- 71
Db 152 DIAFLIDSSGSGINQDFQAKMDPVALMGEPASTSLFSLMQYSNIIKTHFTFEFNKIL 211
Qy 72 DREQIRQGLELEOKLPGSGDTYMHGEPFRASEQIYERKQGYRTA-SYIILTLTGELHED 130
Db 212 DPGSLVDPITVQLQ-----GLTYTATGIRTWBELFHSNGSRKSAKKILVITDQGYRD 266

Qy 131 LFFYSE--REANRSRLGAIYVGVKD-FNE-TQLARI-----ADSKDHVFPVNDGFOA 181
Db 267 PLEKSDVLPADKX---GIIRYALGVGDAPQEPALKEKLNITGAPPDQHFVKVGN-FAA 322
Qy 182 LOGIHSILKSKSCIEIILAEPSTICAGSFQVYVYVNGSFRIARND 227
Db 323 LRSIQROLOEK-----IFAIEGTOSRSSSFQHEMSQEGFSALTSID 364
RESULT 8
ITAD_HUMAN STANDARD; PRT; 1162 AA.
AC Q13349; Q15575; Q15576;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Integrin alpha-D precursor (Leukointegrin alpha D) (CD11d) (ADRB2).
GN Name=ITGAD;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=96111956; PubMed=8777714; DOI=10.1016/1074-7613(95)90058-6;
RA Van der Vlieten M., Le Trong H., Wood C.L., Moore P.F., St John T.,
RT Staunton D.E., Gallatin W.M.,
RL "A novel leukointegrin, alpha d beta 2, binds preferentially to ICAM-3."
RN [2]
RP IMMUNITY 3:683-690(1995).
RX SEQUENCE OF 1-235 FROM N.A.
RX MEDLINE=20187620; PubMed=10722744; DOI=10.1074/jbc.275.12.8959;
RA Noci J.D., Johnson A.K., Dillon J.D.;
RT "Structural and functional characterization of the leukocyte integrin gene CD11d. Essential role of Sp1 and Sp3."
RL J. Biol. Chem. 275:8959-8969(2000).
RN [3]
RP SEQUENCE OF 467-537, 571-602, 633-666, 788-834 AND 910-1125 FROM N.A.
RX MEDLINE=96257236; PubMed=8666289; DOI=10.1016/0378-1119(95)00869-1;
RA Wong D.A., Davis E.M., Lebeau M., Springer T.A.;
RT "Cloning and chromosomal localization of a novel gene-encoding a human beta 2-integrin alpha subunit."
RL Gene 171:291-294(1996).
RN [4]
RP INTERACTION WITH VCAM1.
RX MEDLINE=99059842; PubMed=9841932;
RA Grayson M.H., Van der Vlieten M., Sterbinsky S.A., Michael Gallatin W.,
RT Hoffman P.A., Staunton D.E., Bochner B.S.;
RL "alpha2beta2 integrin is expressed on human eosinophils and functions as an alternative ligand for vascular cell adhesion molecule 1 (VCAM-1)."
RN [5]
RP J. Exp. Med. 188:2187-2191(1998).
RX INTERACTION WITH VCAM1.
RX MEDLINE=99370002; PubMed=10438935;
RA Van der Vlieten M., Crowe D.T., Hoeksma D., Vazeux R., Hoffman P.A.,
RT Grayson M.H., Bochner B.S., Gallatin W.M., Staunton D.E.;
RL "The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a binding interface between I domain and VCAM-1."
RN [6]
RX J. Immunol. 163:1984-1990(1999).
RL -1- FUNCTION: Integrin alpha-D/beta-2 is a receptor for ICAM3 and VCAM1. May play a role in the atherosclerotic process such as clearing lipoproteins from plaques and in phagocytosis of blood-borne pathogens, particulate matter, and senescent erythrocytes from the blood.
CC -1- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-D associates with beta-2.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: Expressed moderately on myelomonocytic cell lines and subsets of peripheral blood leukocytes and strongly on

tissue-specialized cells, including macrophages foam cells within
 CC atherosclerotic plaques, and on splenic red pulp macrophages.
 CC -1- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
 CC with I-domains do not undergo protease cleavage.
 CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
 CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -1- SIMILARITY: Contains 1 VWFA domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see [http://www.ebi.ac.uk/announcements/](http://www.ebi.ac.uk/announcements)
 CC or send an email to license@ebi.ac.uk).

DR EMBL: U37028; AAB36547.1; -;
 DR EMBL: U40274; AAB60634.1; -;
 DR EMBL: U40275; AAB60635.1; -;
 DR EMBL: U40276; AAB60636.1; -;
 DR EMBL: U40277; AAB60637.1; -;
 DR EMBL: U40279; AAB60638.1; -;
 DR EMBL: U40278; AAB60638.1; JOINED.
 DR HSSP: AF187881; AAF62875.1; -;
 DR HSSP: P11215; 1BHQ.
 DR Gene: HGNC:6146; ITGAD.
 DR MIM: 602453; -;
 DR GO: GO:0008305; C:integrin complex; TAS.
 DR GO: GO:0016337; P:cell-cell adhesion; NAS.
 DR GO: GO:007160; P:cell-matrix adhesion; NAS.
 DR GO: GO:006955; P:immune response; NAS.
 DR InterPro: IPR000413; Integrin_alpha.
 DR InterPro: IPR02035; VWF_A.
 DR Pfam: PF01839; FG-GAP; 3_
 DR Pfam: PF00357; Integrin_alpha; 1.
 DR Pfam: PF00092; VWA; 1.
 DR PRINTS: PRO1185; INTEGRINA.
 DR PRINTS: PRO0453; VWFADOMAIN.
 DR SMART: SM00191; Int_alpha; 5.
 DR SMART: SM00327; VWA_1.
 DR PROSITE: PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE: PS50234; VWA; 1.
 DR KEGG: Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
 KW Repeat; Signal; Transmembrane.
 FT CHAIN 1 17 Potential.
 FT DOMAIN 18 1162 Integrin_alpha-D.
 FT TRANSMEM 1101 1121 Extracellular (Potential).
 FT DOMAIN 1122 1162 Potential.
 FT REPEAT 32 85 Cytoplasmic (Potential).
 FT REPEAT 86 7 FG-GAP 1.
 FT DOMAIN 150 332 FG-GAP 2.
 FT REPEAT 350 400 FG-GAP 3.
 FT REPEAT 401 452 FG-GAP 4.
 FT REPEAT 454 516 FG-GAP 5.
 FT REPEAT 518 576 FG-GAP 6.
 FT REPEAT 581 633 FG-GAP 7.
 FT CA_BIND 465 473 Potential.
 FT CA_BIND 530 538 Potential.
 FT CA_BIND 593 601 Potential.
 FT SITE 1127 1131 GPFKR motif.
 FT DISULFID 67 74 By similarity.
 FT DISULFID 106 124 By similarity.
 FT DISULFID 655 710 By similarity.
 FT DISULFID 769 775 By similarity.
 FT DISULFID 846 861 By similarity.
 FT DISULFID 994 1018 By similarity.
 FT DISULFID 1023 1028 By similarity.
 FT CARBOHYD 59 59 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 87 87 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 99 99 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 391 391 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 691 691 N-linked (GlcNAc...) (Potential).

FT CARBOHYD 733 733 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 873 873 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 957 957 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1046 1046 N-linked (GlcNAc...) (Potential).
 FT CONFLICT 500 500 Missing (in Ref. 2).
 FT CONFLICT 515 518 GHPW -> ATP (in Ref. 2).
 FT CONFLICT 825 825 L -> V (in Ref. 2).
 FT CONFLICT 984 984 V -> A (in Ref. 2).
 SQ SEQUENCE 1162 AA; 126885 MW; P296A1A35455D77D CRC64;

Query Match 10.2%; Score 156; DB 1; Length 1162;
 Best Local Similarity 24.3%; Pred.No.0.0015;
 Matches 67; Conservative 50; Mismatches 119; Indels 40; Gaps 11;

Qy 8 DGGPAC-YGGFDLYFLIDKSGSV-LHHWNIYFVEQLAHKFTSPQLRMSPIYFTGTT 65
 Db 139 DATPECHQEMDVFPLIDSGSIDQNDPDMKQFVQAVMGFGSTDTLFLMGYS-----N 134
 Qy 66 LMKLTEDREGIRQGLEBELQKLP---CGDTYHGEFERASEQIYYENRCGYTA-SVIT 120
 Db 195 LKIHFTPTQFRTSPQOSLVDPVQLKGLTFATGILTVVTLFHHKNGARKSARKILI 254
 Qy 121 ALTDGELHEDLFYFSREANRSRDLAGIYCVGVKDFENQGLAR-----IADSKDHF 173
 Db 255 VTIDGQYKDPLEYS-D-VIPQAKAGITRYAIVGHAFCGPTARQELNTSSAPPDHVF 313
 Qy 174 PVNDGFQALQGIHSLILKSCIEIIAASPSTICAGSFQVYVGRNGFRHARNVDVLC-- 231
 Db 314 KV-DNFAALGSIQKQLEK-----IYAVEGQSASASSFQHEMGEGSTALTMDGLFLGA 368
 Qy 232 --SFKINDS-----VTLNKPSVEDTYL 253
 Db 369 VGSFWSGCAFLYPPNMSPTFINMSQENVMDRISYL 404

RESULT 9
 Q9BP08 PRELIMINARY; PRT; 1332 AA.
 ID Q9BP08
 AC Q9BP08
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)
 DE Integrin alpha Htl precursor.
 GN Name=HtrGAL;
 OS Halocynthia roretzi (Sea squirt).
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
 OC Stolidobranchia; Pyuridae; Halocynthia.
 OX NCBI_TaxID=7729;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Hemocyte;
 RX MEDLINE=2103187; PubMed=11160215;
 RA Miyazawa S., Azumi K., Nonaka M.;
 RT "Cloning and characterization of integrin alpha subunits from the
 RT solitary ascidian, Halocynthia roretzi.";
 RL J. Immunol. 166:1710-1715 (2001).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
 CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
 DR EMBL: AB048261; BAB21479.1; -;
 DR HSSP: P11215; 1BHQ.
 DR GO: GO:0016021; C:Integral to membrane; IEA.
 DR GO: GO:0008305; C:integrin complex; IEA.
 DR GO: GO:0005515; F:protein binding; IEA.
 DR GO: GO:0007160; P:cell-matrix adhesion; IEA.
 DR GO: GO:0007229; P:integrin-mediated signaling pathway; IEA.
 DR InterPro: IPR000413; Integrin_alpha.
 DR InterPro: IPR02035; VWF_A.
 DR Pfam: PF00092; VWA; 1.
 DR PRINTS: PRO1185; INTEGRINA.
 DR PRINTS: PRO0453; VWFADOMAIN.
 DR SMART: SM00191; Int_alpha; 5.
 DR SMART: SM00327; VWA; 1.
 DR PROSITE: PS50234; VWA; 1.

KW Cell adhesion; Integrin; Signal; Transmembrane.
 FT SIGNAL 1 30 Potential.
 FT CHAIN 31 1332 Integrin alpha H1.
 SQ SEQUENCE 1332 AA; 145851 MW; 0D9108D2B05CFEAE CRC64;

Query Match 10.0%; Score 153.5; DB 2; Length 1332;
 Best Local Similarity 24.1%; Pred. No. 0.0028; Mismatches 81; Indels 75; Gaps 10;
 Matches 63; Conservative 42; Mismatch 81; Indels 75; Gaps 10;

QY 3 GGRREDG-----PACYGFDLYFILDKSGSVLHHMNEIYFVEQLAHKFISPOLRMSFI 57
 DB 188 GNMKESRNTECS--SGVVLFLVDGSSVGKRFVDKVMVKNIT-----232
 QY 58 VFSTRGTTLMKLTEDREQIQLGSELOKVLPG-----GDTYMHGFERAS 102
 DB 233 -----AKLDIGKEIVRGVGVSHYVEGKSIINKKVIITTEISIGFKLIDNFENAV 283
 QY 103 EQIYENKQGR-----ASVIALTDGELHEDLPFYSREANRS 142
 DB 284 DRI---QLOGTYTGRALOKVIRDPDAYIGNKQVLLLTLDGAKNKILP--NNRRL 338
 QY 143 RDLGAIYVCVGKOFNETOLARIA---DSKDFVPVNDGFOALQGIHSILKSCIEILA 199
 DB 339 RNKGIATFPAVGVEYDISELKLASGTDSTDRVFTYTD-FGELDSIVKSLQTEIQSFVLE 397
 QY 200 AEPSTICAGESPQVVGNGF 220
 DB 398 GKGAKTAG--YEHFNGENF 416

RESULT 10
 Q8NFM1
 ID Q8NFM1 PRELIMINARY; PRT; 1626 AA.
 AC Q8NFM1
 DT 01-OCT-2002 (Tremblrel. 22, Created)
 DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
 DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
 DE Alpha 1 type XXII collagen.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Cartilage;
 RA Koch M., Jin W., Ashworth T., Burgess R.E.;
 RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF406780; AA003620.1; -.
 DR HSSP; P18614; IIMP.
 DR Genew; HGNC:22989; COL22A1.
 DR GO; GO:0005737; C:cytoplasm; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR GO; GO:0007155; P:cell adhesion; IEA.
 DR GO; GO:0006817; P:phosphate transport; IEA.
 DR InterPro; IPR008160; C1g_helix.
 DR InterPro; IPR008160; C1g_helix.
 DR InterPro; IPR008985; Cona_like_1ec_g1.
 DR InterPro; IPR003129; TSP_N.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; Pf01391; Collagen; 16.
 DR Pfam; Pf00092; VMA; 1.
 DR PRINTS; PRO0453; VWFADOMAIN.
 DR Prodom; PD000007; C1g_helix; 7.
 DR SMART; SM00210; TSPN_1.
 DR SMART; SM00327; VMA; 1.
 DR PROSITE; PS50234; VWF_A; 1.
 KW Collagen.
 SQ SEQUENCE 1626 AA; 161115 MW; 34C68B3CFD467407 CRC64;

Query Match 9.7%; Score 148.5; DB 2; Length 1626;
 Best Local Similarity 25.4%; Pred. No. 0.0091;
 Matches 65; Conservative 45; Mismatches 105; Indels 41; Gaps 14;

QY 9 GGPACYG-----GDTLYFILDKSGSV-LHHMNEIYFVEQLAHKF-ISP-QLRMSFI 57
 DB 21 GGGGCCQARAGCKSVHDLVFLDLTSSVGKEDEFEKROWANLVDFEEVGPDRTRGVV 80
 QY 58 VFSTRGTTLMK--LTEDREQIQLGSELOKVLPGGDTYMHGFERASEQIYENKQ-- 112
 DB 81 RYSDRPTTAFELGPGSQEYKAAARL--AYHGGNTGTALRYTTARSPFPAAGRPR 138
 QY 113 ---YRTASVIALTDGELHEDLPFYSREANRSRLGAIYVCVGKDFNETOLARIAD-- 167
 DB 139 DRAK--QVALILTDGR-SQDLVLDAABAAHRA---GIRFPAVGSALEKELEIASP 192
 QY 168 SKDFVPVNDGFOALQGIHSILKSCIEILAEPTICAGESPQVVGNGFPAHNV 227
 DB 193 KSAHVFHVSD-FNAIDIRKLRRLCENVLC--PS-----YAVEGDRFHTNGT 240
 QY 228 RVLCSFKINDSVTLNE 243
 DB 241 KEITGFDLMDLFSYKE 256

RESULT 11
 Q8T6U5
 ID Q8T6U5 PRELIMINARY; PRT; 441 AA.
 AC Q8T6U5;
 DT 01-JUN-2002 (Tremblrel. 21, Created)
 DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
 DE Proximal thread matrix protein 1 variant a.
 OS Mytilus edulis (Blue mussel).
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Periomorpha; Mytiloidea;
 OC Mytiloidea; Mytilidae; Mytilus.
 NCBI_Taxid=6550;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22313036; PubMed=12425661; DOI=10.1021/bm0235903;
 RA Sun C., Lucas J.M., Walte J.H.;
 RT "Collagen-binding matrix proteins from elastomeric extraorganismic
 RT byssal fibers";
 RL Biomacromolecules 3:1240-1248(2002).
 DR EMBL; AF414454; AA183537.1; -.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; Pf00092; VMA; 2.
 DR PRINTS; PRO0453; VWFADOMAIN.
 DR SMART; SM00327; VMA; 2.
 DR PROSITE; PS50234; VWF_A; 2.
 KW Matrix protein.
 SQ SEQUENCE 441 AA; 47543 MW; 881D8BD36891D2B CRC64;

Query Match 9.6%; Score 148; DB 2; Length 441;
 Best Local Similarity 25.5%; Pred. No. 0.002; Mismatches 77; Indels 26; Gaps 12;
 Matches 51; Conservative 46; Mismatch 77; Indels 26; Gaps 12;

QY 13 CYGFGDLYFILDKSGSV---LHHMNEIYFVEQLAHKF--ISP-LRMSFVSTRGT 65
 DB 236 CAGHADIAFVFDASSSTINANNPNVYGLMKDFMDIYDRFKTGPDDGQFVAVFPADATK 295
 QY 66 LMKLTG--DREQIQLGSELOKVLPG--GDTYMHGFERASEQIYENKQ--YRTASV 118
 DB 296 QFGIKDYSSKAEIKGALD--KTPPSITGTAIGDGENARLRY-FPNRRGGGKEEVOKV 351
 QY 119 IIALTDGEL--HEDLPFYSREANRSRLGAIYVCVGKDFNETOLARIADSDHVPV 175
 DB 352 VILITDQNNNGHKS---PEHSSLRKSGVVALGVGTFKSLSELINIASSEIVF-T 406
 QY 176 NDGFOALQGIHSILKSC 195
 DB 407 TSPFNLSKIMENVVYLACM 426

RESULT 12
 Q8T5C3

ID 08TSC3 PRELIMINARY; PRT; 444 AA.
AC 08TSC3;
RT 01-JUN-2002 (TRENBLREL. 21, Created)
DT 01-JUN-2002 (TRENBLREL. 21, Last sequence update)
DE 01-JUN-2003 (TRENBLREL. 24, Last annotation update)
OS Mytilus edulis (Blue mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
NCBI_TaxID=6550;
RN [1]
RP MEDLINE=22313036; PubMed=12425661; DOI=10.1021/bm0255903;
RA Sun C., Lucas J.M., Walte J.H.;
RT "Collagen-binding matrix proteins from elastomeric extraorganic
RT bysral fibers.";
RL Biomatromolecules 3:1240-1248(2002).
DR EMBL; AY053390; AL17973.1;
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR02035; VWF_A.
DR Pfam; PF00092; VMA; 2.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00327; VMA; 2.
DR PROSITE; PS50234; VWF_A; 2.
KM Matrix protein.
SQ SEQUENCE 444 AA; 47815 MW; D2C605347450C931 CRC64;

Query March 9.6%; Score 148; DB 2; Length 444;
Best Local Similarity 25.5%; Pred. No. 0.002;
Matches 51; Conservative 46; Mismatches 77; Indels 26; Gaps 12;

QY 13 CYGQFDYFLIDKSGSV---LHMNEIYFVEQLAHKF--ISPO-LRMSFVSTGTT 65
DB 239 CAGADIAFVFDASSINANNPNVGLMKDMDIVRFNKTGPDGQFAVTFADATK 298
QY 66 LMKLTE--DREQIQGLELOKVLPG--GDTYMEGFRASQDYENRQO--YRTASV 118
DB 299 QFGKDYSSKAEIKGAD--KTPSIIQTAIGDGLNALFV-FPRNGGHEEVQKV 354
QY 119 IIALTDEEL--HEDLFYSEKARSRLDAIVYCVGV-KDFNFTOLARIDSKHVPV 175
DB 355 VILITDGGNNCHK---PEHSSILKRGVIVVIGTGTFKSELINISSEYVF-T 409
QY 176 NDGFQALQGIHSILKSCSI 195
DB 410 TSSFNKLKSKIMENVKLAQM 429

RESULT 13
ITAM HUMAN STANDARD; PRT; 1152 AA.
AC P11215;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha
DE subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1
DE (Neutrophil adhesion receptor).
GN Name=ITGAM; Synonyms=CD11B, CR3A;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86315033; PubMed=2457584;
RA Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;
RT "The human leukocyte adhesion glycoprotein Mac-1 (complement receptor
RT type 3, CD11b) alpha subunit. Cloning, primary structure, and relation
RT to the Integrin, von Willebrand factor and factor B.";
RL J. Biol. Chem. 263:12403-12411(1988).
RN [2]
RP SEQUENCE FROM N.A.

RX MEDLINE=88190151; PubMed=2833753;
RA Arnout M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;
RT "Molecular cloning of the alpha subunit of human and guinea pig
RT leukocyte adhesion glycoprotein Mo1: chromosomal localization and
RT homology to the alpha subunits of integrins.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=88257215; PubMed=2454931; DOI=10.1083/jcb.106.6.2153;
RA Arnout M.A., Gupta S.K., Pierce M.W., Tenen D.G.;
RT "Amino acid sequence of the alpha subunit of human leukocyte adhesion
RT receptor Mo1 (complement receptor type 3).";
RL J. Cell Biol. 106:2153-2158(1988).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=93123748; PubMed=8419480;
RA Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.P., Tenen D.G.;
RT "Structural analysis of the CD11b gene and phylogenetic analysis of
RT the alpha-integrin gene family demonstrate remarkable conservation of
RT genomic organization and suggest early diversification during
RT evolution.";
RL J. Immunol. 150:480-490(1993).
RN [5]
RP SEQUENCE OF 9-1153 FROM N.A.
RX MEDLINE=89098933; PubMed=2563162;
RA Hickey M.J., Ozols J., Baker D.M., Back A.L.,
RA Roth G.J.;
RT "cDNA sequence for the alpha M subunit of the human neutrophil
RT adherence receptor indicates homology to integrin alpha subunits.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:257-261(1989).
RN [6]
RP SEQUENCE OF 1-9 FROM N.A.
RX MEDLINE=92073318; PubMed=1683702;
RA Shelley C.S., Arnout M.A.;
RT "The promoter of the CD11b gene directs myeloid-specific and
RT developmentally regulated expression.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:10525-10529(1991).
RN [7]
RP SEQUENCE OF 1-9 FROM N.A.
RX TISSUE=Blood;
RC MEDLINE=92144966; PubMed=1346576;
RA Pahl H.L., Rosmarin A.G., Tenen D.G.;
RT "Characterization of the myeloid-specific CD11b promoter.";
RL Blood 79:865-870(1992).
RN [8]
RP SEQUENCE OF 17-31.
RX MEDLINE=87076721; PubMed=3539202; DOI=10.1016/0167-4838(86)90037-3;
RA Pierce M.W., Remold-O'Donnell E., Todd R.F. III, Arnout M.A.;
RT "N-terminal sequence of human leukocyte glycoprotein Mo1: conservation
RT across species and homology to platelet IIb/IIIa.";
RL Biochim. Biophys. Acta 874:368-371(1986).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 148-331.
RX MEDLINE=95171458; PubMed=7867070; DOI=10.1016/0022-8674(95)90517-0;
RA Lee J.O., Rieu P., Arnout M.A., Liddington R.;
RT "Crystal structure of the A domain from the alpha subunit of integrin
RT CD3 (CD11b/CD18).";
RL Cell 80:631-638(1995).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 148-334.
RX MEDLINE=96363671; PubMed=8747460; DOI=10.1016/S0969-2126(01)00271-4;
RA Lee J.O., Bankston L.A., Arnout M.A., Liddington R.C.;
RT "Two conformations of the integrin A-domain (I-domain): a pathway for
RT activation?";
RL Structure 3:1333-1340(1995).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 148-337.
RX MEDLINE=98363659; PubMed=9687375; DOI=10.1016/S0969-2126(98)00093-8;
RA Baldwin E.T., Sarver R.W., Bryant G.L. Jr., Curry K.A.,
RA Batshenko M.B., Finzel B.C., Garlick R.L., Heinrichson R.L.,
RA Horton N.C., Kelley L.L., Mildner A.M., Moon J.B., Mott J.E.,
RA Muchner V.T., Tomich C.S., Watpaugh K.D., Wiley V.H.;
RT "Cation binding to the integrin CD11b I domain and activation model

RT assessment." ;
 RL Structure 6:923-935(1998).
 RN [12]
 RP 3D-STRUCTURE MODELING OF 17-616.
 RX MEDLINE=98226734; PubMed=9560195; DOI=10.1073/pnas.95.9.4870;
 RA Ovig C., Springer T.A.;
 RT "Experimental support for a beta-propeller domain in integrin alpha-
 subunits and a calcium binding site on its lower surface." ;
 RL Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875(1998).
 CC -1- FUNCTION: Integrin alpha-M/beta-2 is implicated in various
 adhesive interactions of monocytes, macrophages and granulocytes
 as well as in mediating the uptake of complement-coated particles.
 It is identical with CR-3, the receptor for the IC3b fragment of
 the third complement component. It probably recognizes the R-G-D
 peptide in C3b. Integrin alpha-M/beta-2 is also a receptor for
 fibrinogen, factor X and ICAM1. It recognizes P1 and P2 peptides
 of fibrinogen gamma chain.
 CC -1- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-M
 associates with beta-2.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: Predominantly expressed in monocytes and
 granulocytes.
 CC -1- DOMAIN: The integrin I-domain (insert) is a VFPA domain. Integrins
 with I-domains do not undergo protease cleavage.
 CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
 CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -1- SIMILARITY: Contains 1 VFPA domain.
 CC -1- DATABASE: NAME=PROV; NOTE=CD guide CD11b entry;
 WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd11b.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
 or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL/ J03925; AAA59544.1; -;
 DR EMBL/ M18044; AAA59491.1; -;
 DR EMBL/ J04145; AAA59803.1; -;
 DR EMBL/ S52227; AAB24821.1; -;
 DR EMBL/ S52152; AAB24821.1; JOINED.
 DR EMBL/ S52153; AAB24821.1; JOINED.
 DR EMBL/ S52154; AAB24821.1; JOINED.
 DR EMBL/ S52155; AAB24821.1; JOINED.
 DR EMBL/ S52157; AAB24821.1; JOINED.
 DR EMBL/ S52159; AAB24821.1; JOINED.
 DR EMBL/ S52161; AAB24821.1; JOINED.
 DR EMBL/ S52164; AAB24821.1; JOINED.
 DR EMBL/ S52165; AAB24821.1; JOINED.
 DR EMBL/ S52167; AAB24821.1; JOINED.
 DR EMBL/ S52169; AAB24821.1; JOINED.
 DR EMBL/ S52170; AAB24821.1; JOINED.
 DR EMBL/ S52173; AAB24821.1; JOINED.
 DR EMBL/ S52180; AAB24821.1; JOINED.
 DR EMBL/ S52181; AAB24821.1; JOINED.
 DR EMBL/ S52184; AAB24821.1; JOINED.
 DR EMBL/ S52189; AAB24821.1; JOINED.
 DR EMBL/ S52191; AAB24821.1; JOINED.
 DR EMBL/ S52203; AAB24821.1; JOINED.
 DR EMBL/ S52212; AAB24821.1; JOINED.
 DR EMBL/ S52213; AAB24821.1; JOINED.
 DR EMBL/ S52216; AAB24821.1; JOINED.
 DR EMBL/ S52219; AAB24821.1; JOINED.
 DR EMBL/ S52220; AAB24821.1; JOINED.
 DR EMBL/ S52221; AAB24821.1; JOINED.
 DR EMBL/ S52222; AAB24821.1; JOINED.
 DR EMBL/ S52226; AAB24821.1; JOINED.
 DR EMBL/ M76724; AAB58410.1; -;
 DR EMBL/ M84477; AAA51960.1; -;

DR PIR; A31108; RWHUB.
 DR PDB; 1ABX; Model; @=17-1152.
 DR PDB; 1BHO; X-ray; 1/2=-.
 DR PDB; 1BHQ; X-ray; 1/2=-.
 DR PDB; 1IDN; X-ray; 1/2=-.
 DR PDB; 1IDO; X-ray; @=140-331.
 DR PDB; 1JLM; X-ray; @=143-334.
 DR PDB; 1MLU; X-ray; A=137-331.
 DR PDB; 1MF7; X-ray; A=144-337.
 DR PDB; 1N92; X-ray; A=140-335.
 DR PDB; 1NA5; X-ray; A=144-345.
 DR Genew; HGNC:6149; ITGAM.
 DR MIM; 120980; -.
 DR GO; GO:0008305; C:integrin complex; TAS.
 DR GO; GO:0007155; P:cell adhesion; TAS.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00357; Integrin_alpha; 1.
 DR Pfam; PF00092; VWA; 1.
 DR PRINTS; PR01185; INTEGRINA.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00191; Int_alpha; 5.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS50234; VFPA; 1.
 DR 3D-structure; Calcium; Cell adhesion; Direct protein sequencing;
 KW Glycoprotein; Integrin; Magnesium; Receptor; Repeat; Signal;
 KW Transmembrane.
 FT SIGNAL 1 16
 FT CHAIN 17 1152 Integrin alpha-M.
 FT
 Query Match 9.5%; Score 145.5; DB 1; Length 1152;
 Best Local Similarity 26.2%; Pred. No. 0.01; 81; Indels 43; Gaps 12;
 Matches 60; Conservative 45; Mismatches 81; Indels 43; Gaps 12;
 Qy 18 DLVFLDKSGSVL-HNNNELYFVEQLAKRTISPOLRMSFVSTRGTTMLKLTED---- 72
 Db 150 DIAFLIDGSGSIIIPDRMRKKEFVST-----VMEQLKSKTFLS-----LMQSESEPRRIH 199
 Qy 73 -----RQIQRLGLELQKVLPGCDTVMHGFERSAQIYYENRGYRTA-SVITALT 123
 Db 200 FTFKEPQNNPNPSLVLPITQL--GRTHATATGIRKVRBELFNTNGARKNAFKILVIT 257
 Qy 124 DGEIHEDLFYYSR--RANNSRDGLATVCGVQVDFVETGLAR-----KDHVFP 174
 Db 258 DGEKFGDPLGEDVTEPADRE---GVIRYIVIGVDARSRSEKROELNTIASKPRDHFVQ 314
 Qy 175 VNDGFQALQGIHSHLKSCEIILAAEPSTICAGESFQVAVVNGGFPH 223
 Db 315 VNN-FEALKTITQNLREK----IPALIGTGTGSSSSFEHMSQEGFSA 358
 RESULT 14
 Q8T5C2 PRELIMINARY; PRT; 453 AA.
 AC Q8T5C2;
 DT 01-JUN-2002 (TREMUREl. 21, Created)
 DT 01-JUN-2002 (TREMUREl. 21, Last sequence update)
 DT 01-JUN-2003 (TREMUREl. 24, Last annotation update)
 DE Proximal thread matrix protein 1.
 OS Mytilus galloprovincialis (Mediterranean mussel).
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
 OC Mytiloidea; Mytilidae; Mytilus.
 NC NCBI_TaxID=29156;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22313036; PubMed=12425661; DOI=10.1021/bm0255903;
 RA Sun C., Lucas J.M., Waite J.H.;
 RT "Collagen-binding matrix proteins from elastomeric extraorganic
 RT byssal fibers." ;
 RL Biomacromolecules 3:1240-1248(2002).
 DR EMBL; AY053391; AAL17974.1; -;

DR HSSP; P20701; IMJN.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF00092; VWA; 2.
DR PRINTS; PRO0453; VWFADOMAIN.
DR SMART; SM00327; VWA; 2.
DR PROSITE; PS50234; VWF_A; 2.
DR Metrix protein.
SQ SEQUENCE 453 AA; 48784 MW; D60497F5C0C51E6D CRC64;

Query March 9.4%; Score 145; DR 2; Length 453;
Best Local Similarity 25.5%; Pred. No. 0.0036;
Matches 51; Conservative 44; Mismatches 79; Indels 26; Gaps 12;

QY 13 CYGGEFDYFINDKSGSV---LHMMNEIYFVEQLAHKF--ISPO-LRMSFVSTGTT 65
DB 248 CAGADIAFVFDASSINANNPNYQLMKNMKQIVDFKNTGPDGQFAVVTADATK 307
QY 66 LMKLTE--DREQIRGEEELQKVLPG--GDTYHEGFERASEQIYENRQG--YRTASV 118
DB 308 QFGKDYSSKADIKGAID---KSPSIIIGQTAIGDLENALLEV-FPNRGGGEEVQKV 363
QY 119 IIALTDEL--HELFYSEAREARSDLAIVYCVG-KDFNTQLARTADSDHPV 175
DB 364 VILLTDDQNGHKS---PEHSSILRKEGVVAIGVGTFPLSELINIASSEYVF-T 418

QY 176 NDGFQALQGIHSILKSCI 195
DB 419 TSSFDKSKIMEDVYKLAQM 438

RESULT 15
ITAX HUMAN STANDARD; PRT; 1163 AA.
AC P20702; O81VA6; 17, Created
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Integrin alpha-X precursor (Leukocyte adhesion glycoprotein p150,95
alpha chain) (Leukocyte adhesion receptor p150,95) (CD11c) (Lew M5).
GN Name=ITGAX; Synonyms=CD11c;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8616645; PubMed=3327687;
RA Corbi A.L., Miller U.J., O'Connor K., Larson R.S., Springer T.A.;
RT "cDNA cloning and complete primary structure of the alpha subunit of a
leukocyte adhesion glycoprotein, p150,95."
RL EMBO J. 6:4023-4028(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90153906; PubMed=2303426;
RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;
RT "Genomic structure of an integrin alpha subunit, the leukocyte p150,95
molecule."
RL J. Biol. Chem. 265:2782-2788(1990).
RN [3]
RP ERRATUM.
RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;
RL J. Biol. Chem. 265:12750-12751(1990).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUEBiood;
RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Aleschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heleh F.,
RA Datschenko L., Marusina K., Farmer A.A., Rudin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren R.J., Lu X., Gibbs R.A.,
RA Rahy J., Helton R., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP SEQUENCE OF 20-43.
RX MEDLINE=87167596; PubMed=3549901;
RA Miller L.J., Wiebe M., Springer T.A.;
RT "Purification and alpha subunit N-terminal sequences of human Mac-1
and p150,95 leukocyte adhesion proteins."
RL J. Immunol. 138:2381-2383(1987).
CC -1- FUNCTION: Integrin alpha-X/beta-2 is a receptor for fibrinogen. It
recognizes the sequence G-P-R in fibrinogen. It mediates cell-cell
interaction during inflammatory responses. It is especially
important in monocyte adhesion and chemotaxis.
CC -1- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-X
associates with beta-2.
CC -1- TISSUE SPECIFICITY: Predominantly expressed in monocytes and
granulocytes.
CC -1- DOMAIN: The integrin I-domain (inset) is a VWA domain. Integrins
with I-domains do not undergo protease cleavage.
CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
CC -1- SIMILARITY: Contains 1 VWA domain.
CC -1- DATABASE: NAME=PROX; NOTE=CD guide CD11c entry;
WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd11c.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M8165; AA59180.1; -;
DR EMBL; M29165; -; NOT ANNOTATED CDS.
DR EMBL; M29487; AA51620.1; ALT_SEQ.
DR EMBL; M29482; AA51620.1; JOINED.
DR EMBL; M29483; AA51620.1; JOINED.
DR EMBL; M29484; AA51620.1; JOINED.
DR EMBL; M29485; AA51620.1; JOINED.
DR EMBL; M29486; AA51620.1; JOINED.
DR EMBL; BC038237; AA938237.1; -;
DR PIR; A36584; RWHUIC.
DR PDB; 1N3Y; X-ray; A=141-338.
DR Genew; HGNC:6152; ITGAX.
DR MIM; 151510; -;
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0004872; F:receptor activity; TAS.
DR GO; GO:0007155; P:cell adhesion; TAS.
DR GO; GO:0009887; P:organogenesis; TAS.
DR InterPro; IPR00413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PRO1839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_alpha; 1.
DR Pfam; PF00092; VWA; 1.
DR PRINTS; PRO1185; VWFADOMAIN.
DR PRINTS; PRO0453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.


```

FT      HELIX      326      334
FT      TURN      335      335
SO      SEQUENCE   1163 AA; 127886 MW; 83658A13B5C5DE8F CRC64;

Query Match
Query local Similarity 24.8%; Score 143.5; DB 1; Length 1163;
Best local Similarity 24.8%; Pred. No. 0.015;
Matches 55; Conservative 43; Mismatches 89; Indels 35; Gaps 10;

QY      18  DLYPIIDKSGSV-LHMHNEIYYFVEQLAHKFIISPOLMMSFIVFSTRGTTLMKLTEDRE-- 74
       151  DIVFLIDGSGSISISRNPFATMMNFVRAVISOFRPSTQFSIMQFANKFQTHFTFEEFRRTS 210
       75  ---QIRGGLBELQKVLPGSDPTVMHSEGERASEQIYYENRGQYRTAS-VIATLTDEGLHED 130
       211  NPLSLASVHQLQ----GFTYTATAIQNVAVHRLFHASYGARRDATKILIVITDGKKEGD 265
QY      131  LFFYSEREARSRDLGAIIVYCVGV-----KDFNETQLARIAD--SKDHVFPVNDG 178
       266  SLDYKD-VIEMADAGIIRYAIGGLAFQNRNSMKEIND-----IASKPSQEHIFXVED- 318
QY      179  FOALQGIHSHILKKSCTEIIAEPSTICAGESFQVVVRGNGF 220
       319  FDALKDIOQLKEK----IPAIEGTETFTSSSFELMAQEGF 356
Db

```

THIS PAGE BLANK (USPTO)


```

61 CTGGTCTCATCTGCGCGCGGCGAAGGAGGAGATGAGGAGTCCAGCTGCTAC 120
DB CTGGTCTCATCTGCGCGCGGCGAAGGAGGAGATGAGGAGTCCAGCTGCTAC 223
QY 121 GCGGAGTTTGAAGCTGTAATCTTTTGAAGAAATGAGAGTGTCTGACCACTGGAAAT 180
DB 224 GCGGAGTTTGAAGCTGTAATCTTTTGAAGAAATGAGAGTGTCTGACCACTGGAAAT 283
QY 181 GAAATCTAATCTTGTGGAGAGTTGGCTCAAAATTCATGAGCCCAAGTTGAGATG 240
DB 284 GAAATCTAATCTTGTGGAGAGTTGGCTCAAAATTCATGAGCCCAAGTTGAGATG 343
QY 241 TCCCTTAATGTTTCTCCACCCGAGGAACTTAATGAAATCTGACGAAAGACAGAGA 300
DB 344 TCCCTTAATGTTTCTCCACCCGAGGAACTTAATGAAATCTGACGAAAGACAGAGA 403
QY 301 CAAATCCGTCAAGGCGCTGAGAGAACTCCAGAAAGTTTCCGAGGAGAGACCTTACATG 360
DB 404 CAAATCCGTCAAGGCGCTGAGAGAACTCCAGAAAGTTTCCGAGGAGAGACCTTACATG 463
QY 361 CATGAAGATTTGAAGGCGCTGAGAGAGATTTATTAAGAAACAGACAAAGGATACAG 420
DB 464 CATGAAGATTTGAAGGCGCTGAGAGAGATTTATTAAGAAACAGACAAAGGATACAG 523
QY 421 ACAGCCAGCGTCAATGCTTGAAGTCTGATGAGAGAACTCCAGAAAGTCTCTTTTCTAT 480
DB 524 ACAGCCAGCGTCAATGCTTGAAGTCTGATGAGAGAACTCCAGAAAGTCTCTTTTCTAT 583
QY 481 TCAGAGAGGAGGCTTAATGCTTGAAGATCTTGTGCAATTTGTTTACTGTGTGTGTG 540
DB 584 TCAGAGAGGAGGCTTAATGCTTGAAGATCTTGTGCAATTTGTTTACTGTGTGTGTG 643
QY 541 AAAAGATTTCAATGAGACACAGCTGCGCGGATTTGGGAGACATGATGATGTTTCTCC 600
DB 644 AAAAGATTTCAATGAGACACAGCTGCGCGGATTTGGGAGACATGATGATGTTTCTCC 703
QY 601 GTGAATGACGCGCTTCAAGCTCTGCAAGGCACTCACTCAATTTTGAAGAGTCTGCG 660
DB 704 GTGAATGACGCGCTTCAAGCTCTGCAAGGCACTCACTCAATTTTGAAGAGTCTGCG 763
QY 661 ATCGAAATTTCTGAGCTGAGACCACTCACTCAATGTCAGAGAGTCACTTTCAATGTTCT 720
DB 764 ATCGAAATTTCTGAGCTGAGACCACTCACTCAATGTCAGAGAGTCACTTTCAATGTTCT 823
QY 721 GTGAGAGGAAAGCGCTTCCGACATGCGCGCACTGAGACAGGAGTCTCTGAGCTTCAAG 780
DB 824 GTGAGAGGAAAGCGCTTCCGACATGCGCGCACTGAGACAGGAGTCTCTGAGCTTCAAG 883
QY 781 ATCAAATGACTCGGTCACTCAATGAGAGCCCTTTTCTGTGAAAGACACTTATTTACTG 840
DB 884 ATCAAATGACTCGGTCACTCAATGAGAGCCCTTTTCTGTGAAAGACACTTATTTACTG 943
QY 841 TGTCCAGGCGCTTATCTTAAAGAGTGGCATGAAAGTGCATCTCAAGTCAAGTGAAC 900
DB 944 TGTCCAGGCGCTTATCTTAAAGAGTGGCATGAAAGTGCATCTCAAGTCAAGTGAAC 1003
QY 901 GATGCGCTCTCTTTATCTCCAGTCTGTCACTATCAACACACACTGTTCTGAGGAT 960
DB 1004 GATGCGCTCTCTTTATCTCCAGTCTGTCACTATCAACACACACTGTTCTGAGGAT 1063
QY 961 TCCATCTGTGCGCATGCGCGCTGTGATCTGTTCTGTGCTTGAAGCCCTGCGCTCTCTG 1020
DB 1064 TCCATCTGTGCGCATGCGCGCTGTGATCTGTTCTGTGCTTGAAGCCCTGCGCTCTCTG 1123
QY 1021 TGGTCTGTGCGCTCTGTGCTGACTGTGATTAACAGAGAGTCCCTTCCACCCCTGCGAG 1080
DB 1124 TGGTCTGTGCGCTCTGTGCTGACTGTGATTAACAGAGAGTCCCTTCCACCCCTGCGAG 1183
QY 1081 GAGAGTGAAGAAATTAATTAATA 1104
DB 1184 GAGAGTGAAGAAATTAATTAATA 1207

```

```

RESULT 2
AF421380
LOCUS 1414 bp mRNA linear PRI 13-NOV-2001
DEFINITION Homo sapiens anthrax toxin receptor mRNA, complete cds.
ACCESSION AF421380
VERSION AF421380.1 GI:16566412
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Bradley,K.A., Mougidge,J., Mougidge,M., Collier,R.J. and Young,J.A.
TITLE Identification of the cellular receptor for anthrax toxin
JOURNALS Nature 414 (6860), 225-229 (2001)
MEDLINE 21557240
PUBMED 11700562
REMARK http://www.nature.com
2 (bases 1 to 1414)
REFERENCE Bradley,K.A., Mougidge,J., Mougidge,M., Collier,R.J. and Young,J.A.
AUTHORS Direct Submission
TITLE Submitted (19-SEP-2001) Department of Oncology, University of
JOURNALS Wisconsin-Madison, 1400 University Ave., Madison, WI 53706, USA
FEATURES
source
1..1414
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="Hela"
1..103
104..1210
104..1210
/codon_start=1
/product="anthrax toxin receptor"
/protein_id="P4126486.1"
/db_xref="GI:16566413"
/translation="MATERRRALGIGFQWLSLATVLICAGGGRREDGPGAGTGD
LPLFLDSGSLVHHMNEIYFVQLAKFLSPOLRMSPIVSTGTLMLKLTEDREI
ROGIELEOKVLPQGDITMHEGFERASEQIYVENOGRTASVIALTDGSEILKPY
SEREANSRLGAIIVCVGKDFNEIOLAIAIDSKHVPVNDGFOLGDIHSILK
SCIEPIIAERSTICAGESPOVYVRNGNFRHARIVRLCSFKINDSVLNEKPSVD
TYLICPAIPILKEVGMKRALQVSMKDGSLISSVITITTHCSDSIIAIALILFLLL
ALALWFWPLCCVILIKVPPAESSENKIK"
1211..1414
ORIGIN
3'UTR
Query Match 100.0%; Score 1104; DB 9; Length 1414;
Blast Local Similarity 100.0%; Pident. No. 0;
Matches 1104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCCACGCGGAGGAGAGAGCCCTCGGCATCGGCTTCCAGTGGCTCTCTTTGGCCACT 60
DB 104 ATGGCCACGCGGAGGAGAGAGCCCTCGGCATCGGCTTCCAGTGGCTCTCTTTGGCCACT 163
QY 61 CTGGTCTCATCTGCGCGCGGCGAAGGAGGAGAGATGAGGAGTCCAGCTGCTAC 120
DB 164 CTGGTCTCATCTGCGCGCGGCGAAGGAGGAGAGATGAGGAGTCCAGCTGCTAC 223
QY 121 GCGGAGTTTGAAGCTGTAATCTTTTGAAGAAATGAGAGTGTCTGACCACTGGAAAT 180
DB 224 GCGGAGTTTGAAGCTGTAATCTTTTGAAGAAATGAGAGTGTCTGACCACTGGAAAT 283
QY 181 GAAATCTAATCTTGTGGAGAGTTGGCTCAAAATTCATGAGCCCAAGTTGAGATG 240
DB 284 GAAATCTAATCTTGTGGAGAGTTGGCTCAAAATTCATGAGCCCAAGTTGAGATG 343
QY 241 TCCCTTAATGTTTCTCCACCCGAGGAACTTAATGAAATCTGACGAAAGACAGAGA 300
DB 344 TCCCTTAATGTTTCTCCACCCGAGGAACTTAATGAAATCTGACGAAAGACAGAGA 403
QY 301 CAAATCCGTCAAGGCGCTGAGAGAACTCCAGAAAGTCTGAGAGAGAGACCTTACATG 360

```

Db 404 CAATCCGTCAGAGGCTTGAAGAACTCCAGAAAGTTCTGCGAGAGAGACACTTACATG 463
 Qy 361 CATGAAGATTTGAAAGGGCCAGTGAAGCATTTATTAAGAAACAGACAGAGTACAG 420
 Db 464 CATGAAGATTTGAAAGGGCCAGTGAAGCATTTATTAAGAAACAGACAGAGTACAG 523
 Qy 421 AAGCCAGGTCATCTTCTTGTGATGAGAACTCCATGAAGATCTTTTTCAT 480
 Db 524 AAGCCAGGTCATCTTCTTGTGATGAGAACTCCATGAAGATCTTTTTCAT 583
 Qy 481 TCAGAGAGGAGGCTTAATAGTCTGAGATCTTGTGCAATTTACTGTGTGTG 540
 Db 584 TCAGAGAGGAGGCTTAATAGTCTGAGATCTTGTGCAATTTACTGTGTGTG 643
 Qy 541 AAAAGATTTCAATGAGACACAGCTGGCCCGAATTCGAGACATGAAGTATGTTTCCC 600
 Db 644 AAAAGATTTCAATGAGACACAGCTGGCCCGAATTCGAGACATGAAGTATGTTTCCC 703
 Qy 601 GTGAATGAGGCTTTCAAGGCTCTGCAAGGATCATCTCAATTTTGAAGATCTGCG 660
 Db 704 GTGAATGAGGCTTTCAAGGCTCTGCAAGGATCATCTCAATTTTGAAGATCTGCG 763
 Qy 661 ATCGAAATTTAGACAGCTGAACCATCAATATGTCAGAGAGTCAATTTGAAGTTCG 720
 Db 764 ATCGAAATTTAGACAGCTGAACCATCAATATGTCAGAGAGTCAATTTGAAGTTCG 823
 Qy 721 GTGAGAGAAACGGCTTCCGACATGCGCCGAACTGAGACAGGCTCTTGCAGCTTCAAG 780
 Db 824 GTGAGAGAAACGGCTTCCGACATGCGCCGAACTGAGACAGGCTCTTGCAGCTTCAAG 883
 Qy 781 ATCAATGACTCGGTCACTCAATGAAGAGCCCTTTCTGTGAAGACATTTATTTACTG 840
 Db 884 ATCAATGACTCGGTCACTCAATGAAGAGCCCTTTCTGTGAAGACATTTATTTACTG 943
 Qy 841 TGTCCAGGCTCTATCTTAAAGAGTGGCATGAAGAGTGCATCCAGTCCAGTCAAGTGAAC 900
 Db 944 TGTCCAGGCTCTATCTTAAAGAGTGGCATGAAGAGTGCATCCAGTCCAGTCAAGTGAAC 1003
 Qy 901 GATGCGCTCTCTTTATCTCAAGTCTGTGATCATCAACACACACACACTGTTCTGAGGT 960
 Db 1004 GATGCGCTCTCTTTATCTCAAGTCTGTGATCATCAACACACACACACTGTTCTGAGGT 1063
 Qy 961 TCCATCTGCGCATGCGCTGCTGATCTGTCTGCTCTGAGCCCTGCTCTCTGAG 1020
 Db 1064 TCCATCTGCGCATGCGCTGCTGATCTGTCTGCTCTGAGCCCTGCTCTCTGAG 1123
 Qy 1021 TGGTCTGCGCCCTCTGCTGCACTGTGATTAACAAGAGTCCCTCCACCCCTGCGAG 1080
 Db 1124 TGGTCTGCGCCCTCTGCTGCACTGTGATTAACAAGAGTCCCTCCACCCCTGCGAG 1183
 Qy 1081 GAGAGTGAGAAATTAATAATAA 1104
 Db 1184 GAGAGTGAGAAATTAATAATAA 1207

RESULT 3
 AX393246 5540 bp DNA linear PAT 23-MAR-2002
 LOCUS AX393246
 DEFINITION Sequence 176 from Patent W00210217.
 ACCESSION AX393246
 VERSION AX393246.1 GI:19701296
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE
 1 St Croix, B., Kinzler, K.W. and Vogelstein, B.
 Endothelial cell expression patterns
 Patent: WO 0210217-A 176 07-FEB-2002;
 The Johns Hopkins University (US)
 FEATURES
 source location/Qualifiers
 1..5540

/organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 98.8%; Score 1090.8; DB 6; Length 5540;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1092; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGCCACGGGCGAGAGAGAGCCCTCGGCAATCGGCTTCCAGTGGCTCTTTTGGCACT 60
 Db 144 ATGCCACGGGCGAGAGAGCCCTCGGCAATCGGCTTCCAGTGGCTCTTTTGGCACT 203
 Qy 61 CTGGTCTCATCTGCGCGGCGAAGGGGAGAGAGAGATGGGGGTCCAGCCCTTAC 120
 Db 204 CTGGTCTCATCTGCGCGGCGAAGGGGAGAGAGATGGGGGTCCAGCCCTTAC 263
 Qy 121 GCGGATTTGACCTGTACTTCAATTTTGAACAATAGAGATGTGCTGACCATCTGAAT 180
 Db 264 GCGGATTTGACCTGTACTTCAATTTTGAACAATAGAGATGTGCTGACCATCTGAAT 323
 Qy 181 GAATCTATTAATTTGTGAAACGTTGGCTCAAAATTCATCAGCCCAAGTGAAGATG 240
 Db 324 GAATCTATTAATTTGTGAAACGTTGGCTCAAAATTCATCAGCCCAAGTGAAGATG 383
 Qy 241 TCTTTATTTGTTTCTCCACCCGAGAACCACTTAATGAAGCTGACAGAAACAGAGAA 300
 Db 384 TCTTTATTTGTTTCTCCACCCGAGAACCACTTAATGAAGCTGACAGAAACAGAGAA 443
 Qy 301 CAATCCGTCAAGGCTTGAAGAACTCCAGAAAGTTCTGCGAGAGAGACATTTACATG 360
 Db 444 CAATCCGTCAAGGCTTGAAGAACTCCAGAAAGTTCTGCGAGAGAGACATTTACATG 503
 Qy 361 CATGAAGATTTGAAGGGCCAGTGAAGATTTATTAAGAAACAGACAGAGGTACAG 420
 Db 504 CATGAAGATTTGAAGGGCCAGTGAAGATTTATTAAGAAACAGACAGAGGTACAG 563
 Qy 421 AAGCCAGGCTCATCTTGTGATCTGATGAGAACTCCATGAAGATCTCTTTTCTAT 480
 Db 564 AAGCCAGGCTCATCTTGTGATCTGATGAGAACTCCATGAAGATCTCTTTTCTAT 623
 Qy 481 TCAAGAGAGAGGCTTAATAGTCTGAGATCTTGGTCAATTTTACTGTGTGTG 540
 Db 624 TCAAGAGAGAGGCTTAATAGTCTGAGATCTTGGTCAATTTTACTGTGTGTG 683
 Qy 541 AAAAGATTTCAATGAGACACAGCTGGCCCGAATTCGAGACATGAAGATCATGTTC 600
 Db 684 AAAAGATTTCAATGAGACACAGCTGGCCCGAATTCGAGACATGAAGATCATGTTC 743
 Qy 601 GTGAATGAGGCTTTCAAGGCTCTGCAAGGATCATCTCAATTTTGAAGAGTCTGCG 660
 Db 744 GTGAATGAGGCTTTCAAGGCTCTGCAAGGATCATCTCAATTTTGAAGAGTCTGCG 803
 Qy 721 GTGAGAGAAACGGCTTCCGACATGCGCCGAACTGAGACAGGCTCTTGCAGCTTCAAG 780
 Db 864 GTGAGAGAAACGGCTTCCGACATGCGCCGAACTGAGACAGGCTCTTGCAGCTTCAAG 923
 Qy 781 ATCAATGACTCGGTCACTCAATGAAGAGCCCTTTCTGTGAAGATCTTATTTACTG 840
 Db 924 ATCAATGACTCGGTCACTCAATGAAGAGCCCTTTCTGTGAAGATCTTATTTACTG 983
 Qy 841 TGTCCAGGCTCTATCTTAAAGAGTGGCATGAAGCTGACCTCCAGTCCAGTCAAGTGAAC 900
 Db 984 TGTCCAGGCTCTATCTTAAAGAGTGGCATGAAGCTGACCTCCAGTCCAGTCAAGTGAAC 1043
 Qy 901 GATGCGCTCTCTTTATCTCAAGTCTGTGATTAACAACACACACACTGTTCTGAGGT 960
 Db 1044 GATGCGCTCTCTTTATCTCAAGTCTGTGATTAACAACACACACACTGTTCTGAGGT 1103

QY	961	TCATATCCCTGGGCATATGGCCCTGCTATATCCGTTCCTGCTCCATGACCCTGGATCTCTCTGG	1020
Db	1104	TCATATCCCTGGGCATATGGCCCTGCTATATCCGTTCCTGCTCCATGACCCTGGATCTCTCTGG	1165
QY	1021	TGATTCTTGCCCTCTGCTGCACTGTGTATATCAAGAGAGTTCCTCCACCCCTGCGCAG	1080
Db	1164	TGGTTCTTGCCCTCTGCTGCACTGTGTATATCAAGAGAGTTCCTCCACCCCTGCGCAG	1223
QY	1081	GAGAGTGAAGAAA	1094
Db	1224	GAGAGTGAAGAAA	1237
RESULT 4			
AX393301		5540 bp	DNA linear PAT 23-MAR-2002
LOCUS	AX393301		
DEFINITION	Sequence 231 from Patent WO2010217.		
ACCESSION	AX393301		
VERSION	AX393301.1	GI:19701322	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE	1	St Croix B, Kinzler K.W. and Vogelstein B.	
AUTHORS	Endothelial cell expression patterns		
TITLE	Patent: WO 0210217-A 231 07-FEB-2002;		
JOURNAL	The Johns Hopkins University (US)		
FEATURES	Location/Qualifiers		
source	1..5540		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	98.8%;	Score 1090.8;	DB 6; Length 5540;
Best Local Similarity	99.8%;	Pred. No. 0;	
Matches 1092;	Conservative	0; Mismatches	2; Indels
			Gaps 0;
QY	1	ATGGCCACGGCGGAGCGGAGAGACCTTCGGCATCGGCTTCCAGTGGCTCTTTTGGCACT	60
Db	144	ATGGCCACGGCGGAGCGGAGAGACCTTCGGCATCGGCTTCCAGTGGCTCTTTTGGCACT	203
QY	61	CTGGTGTCTCATCTGGCCGGGCAAGGGGAGCGCAGGAGAGATGGGGGTCCAGCTGTCTAC	120
Db	204	CTGGTGTCTCATCTGGCCGGGCAAGGGGAGCGCAGGAGAGATGGGGGTCCAGCTGTCTAC	263
QY	121	GGCGGATTGGACCTGTACTCTTATTTTGGCAAAATCAGGAAGTGTCTGCAACACTGGAAT	180
Db	264	GGCGGATTGGACCTGTACTCTTATTTTGGCAAAATCAGGAAGTGTCTGCAACACTGGAAT	323
QY	181	GAAATCTATTACTTTTGTGAAACAGTTGGCTCAAAATTCATCAGCCACAGTTGGAATG	240
Db	324	GAAATCTATTACTTTTGTGAAACAGTTGGCTCAAAATTCATCAGCCACAGTTGGAATG	383
QY	241	TCCTTTATTTGTTTCTCCACCCGAGAACACTTATGAAACTGACAGAAACAGAGAA	300
Db	384	TCCTTTATTTGTTTCTCCACCCGAGAACACTTATGAAACTGACAGAAACAGAGAA	443
QY	301	CAAAATCCGTCMAAGGCTGAGAAACATCCCAAGAAAGTTCTGCCAGAGAGACACTTAACATG	360
Db	444	CAAAATCCGTCMAAGGCTGAGAAACATCCCAAGAAAGTTCTGCCAGAGAGACACTTAACATG	503
QY	361	CATGAAGATTTGAAAGAGGCCCAGTAGAGCAGATTTATATGAAAAAGACAAAGGATACAGG	420
Db	504	CATGAAGATTTGAAAGAGGCCCAGTAGAGCAGATTTATATGAAAAAGACAAAGGATACAGG	563
QY	421	ACAGCCAGCGTCATCATCTTTGACTGTGAACTGATGAGAACTTCATGAAGATCTCTTTTCTAT	480
Db	564	ACAGCCAGCGTCATCATCTTTGACTGTGAACTGATGAGAACTTCATGAAGATCTCTTTTCTAT	623
QY	481	TCAGAGAGGAGCGTATATAGTCTCGAAGATCTTTGGTCAATCTTTACTCTGTGTGTGTG	540

Db	624	TCAGAGGGAGGCTAATGAGTCTCGAGATCTTGAGTCAATTTGTTACTGTGTGTGTCG	683
Oy	541	AAAGATTTCAATAGACACAGCTGGCCCCGGAATGGCGGACAGTAAAGATCAATGTGTCCC	600
Db	684	AAAGATTTCAATAGACACAGCTGGCCCCGGAATGGCGGACAGTAAAGATCAATGTGTCCC	743
Oy	601	GTGAATGACGGCTTTCAAGGCTGTGCAGAGGACATCCACTCAATTTTGAAGAACTCTGC	660
Db	744	GTGAATGACGGCTTTCAAGGCTGTGCAGAGGACATCCACTCAATTTTGAAGAACTCTGC	803
Oy	661	ATCGAATTTTAGAGCTGAACCATCCACATATGTGCAGAGAGTGATTTCAAGTTGTC	720
Db	804	ATCGAATTTTAGAGCTGAACCATCCACATATGTGCAGAGAGTGATTTCAAGTTGTC	863
Oy	721	GTGAGAGAAACGGCTTCCGACATGCCCGGACAGTGAACAAGGTCCTCTGCAGCTTCAAG	780
Db	864	GTGAGAGAAACGGCTTCCGACATGCCCGGACAGTGAACAAGGTCCTCTGCAGCTTCAAG	923
Oy	781	ATCATATGACTGGTGCACACTGCATGAGAGCCCTTTTCTGTGAAAGACATTATTTACTG	840
Db	924	ATCATATGACTGGTGCACACTGCATGAGAGCCCTTTTCTGTGAAAGATTAATTATTTACTG	983
Oy	841	TGTCACAGCGCTATCTTAAAGAAAGTTGGCATGAAGACTGCATCCAGAGTCCAGATCAAC	900
Db	984	TGTCACAGCGCTATCTTAAAGAAAGTTGGCATGAAGACTGCATCCAGAGTCCAGATCAAC	1043
Oy	901	GATGGCCTCTCTTTTATCTCCAGTTCTGTCTATCATCAACACACACACACTGTTCTGAAGGT	960
Db	1044	GATGGCCTCTCTTTTATCTCCAGTTCTGTCTATCATCAACACACACACTGTTCTGAAGGT	1103
Oy	961	TCCATCTCGGCAATCGGCTGTGTGATCTGTGTCCTGACTGTAAGGCTGTGCTCTCTGAG	1020
Db	1104	TCCATCTCGGCAATCGGCTGTGTGATCTGTGTCCTGACTGTAAGGCTGTGCTCTCTGAG	1163
Oy	1021	TGTTTCTGGCCCCCTCTGCTGCATGTGATTAATCAAGAGAGTCCCTCCACCCCTGCGGAG	1080
Db	1164	TGTTTCTGGCCCCCTCTGCTGCATGTGATTAATCAAGAGAGTCCCTCCACCCCTGCGGAG	1223
Oy	1081	GAGAGTGAAGAAA 1094	
Db	1224	GAGAGTGAAGAGA 1237	

LOCUS	AX458370	5540 bp	DNA	linear	PAT 08-JUL-2002
DEFINITION	Sequence 5 from Patent WO0246228.				
ACCESSION	AX458370				
VERSION	AX458370.1	GI:21725040			
KEYWORDS	.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1				
AUTHORS	Young, J.A., Bradley, K.A., Collier, R.J. and Moggridge, J.S.				
TITLE	Receptor for B. Anthracis toxin				
JOURNAL	Parent: WO 0246228-A 5 13 -JUN-2002;				
	WISCONSIN ALUMNI RESEARCH FOUNDATION (US)				
FEATURES	Location/Qualifiers				
source	1..5540				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				
	144..1838				
	/note="unnamed protein product"				
	/codon_start=1				
	/protein_id="CAD38350.1"				
	/db_xref="GI:21725041"				
	/translation="MAAARARALGIGROWLSTALVLICAGGGRRRGGACVGGF				
	LYFLDSSGVLAHMMNFIYFVEQLAKRFSPLQRMISFGVSTGTILMKLTEDREQ				
	RQGEELQKVLPGGDITVHRRGFEASQIYYENQGRKVASVIALTDGELHEDLFF				

```

/notice="unnamed protein product"
/codon_start=1
/protein_id="CAD38350.1"
/db_xref="GI:21725011"
/translation="MTAERSLALGIFQWLSLALYLITAGGGRRERDGGPACTYGGF
LYFLITLAKVLLHMEIETLYEVQGLHKRTSPLQNLRSFLVSTRTGLTKRLDEKQ
DQGGLETLQKVLPGDDYIMEGFRASEQSLTYENRCQRYASVIALTGGLEHEDLPF

```


sig_peptide 144..224
/gene="TEM8"
mac_peptide 225..1835
/gene="TEM8"
/product="tumor endothelial marker 8"
misc_feature 1104..1172
/gene="TEM8"
/note="Region: predicted transmembrane domain"

ORIGIN

Query Match 98.8%; Score 1090.8; DB 9; Length 5540;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1092; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGGCCAGCGCGGAGAGAGCCCTCGGCATCGGCTTCCAGTGGCTCTCTTTGGCCACT 60
144 ATGGCCAGCGCGGAGAGAGCCCTCGGCATCGGCTTCCAGTGGCTCTCTTTGGCCACT 203
61 CTGGTCTCATCTGCGCGGCGGAGAGAGGAGATGGGGTTCAGCTGTAC 120
204 CTGGTCTCATCTGCGCGGCGGAGAGAGGAGATGGGGTTCAGCTGTAC 263
121 GCGGATTTGACCTTACTCTTATTTTGAAGAAATGAGAGAGTGTGACCACTGGAGAT 180
264 GCGGATTTGACCTTACTCTTATTTTGAAGAAATGAGAGAGTGTGACCACTGGAGAT 323
181 GAAATCTATTAATTGTGGAGAGATGGCTCAAAATTCATGAGCCCAAGTTGAGAAATG 240
324 GAAATCTATTAATTGTGGAGAGATGGCTCAAAATTCATGAGCCCAAGTTGAGAAATG 383
241 TCCTTTATTTGTTTCTCCACCCGAGAGAACCTTTAATGAACGTGACAGAGACAGAGA 300
384 TCCTTTATTTGTTTCTCCACCCGAGAGAACCTTTAATGAACGTGACAGAGACAGAGA 443
301 CAAATCCGTCAGAGGCTGAGAGAACTCCAGAAAGTTCTGCGAGAGAGACCTTACATG 360
444 CAAATCCGTCAGAGGCTGAGAGAACTCCAGAAAGTTCTGCGAGAGAGACCTTACATG 503
361 CATGAAGATTTGAAGAGGCGCAGTAGAGAGATTTTATGAAGAACAGAGGATACAGG 420
504 CATGAAGATTTGAAGAGGCGCAGTAGAGAGATTTTATGAAGAACAGAGGATACAGG 563
421 ACAGCAGGCTCATCTGCTTTGACTGATGAGAGAACTCCATGAAGATCTCTTTTCTAT 480
564 ACAGCAGGCTCATCTGCTTTGACTGATGAGAGAACTCCATGAAGATCTCTTTTCTAT 623
481 TCAAGAGAGGAGGCTAATAGTCTGAGAGATCTTGGTGAATGTTTACTGCTGTGTG 540
624 TCAAGAGAGGAGGCTAATAGTCTGAGAGATCTTGGTGAATGTTTACTGCTGTGTG 683
541 AAAGATTTCAATGAGACACAGCTGGCCCGAATTTGCGAGAGATGAGATCATGTGTTCC 600
684 AAAGATTTCAATGAGACACAGCTGGCCCGAATTTGCGAGAGATGAGATCATGTGTTCC 743
601 GTGAATGACGCTTTTCAGGCTCTGCAAGGATCATCTCAATTTTGAAGAGTCTGCG 660
744 GTGAATGACGCTTTTCAGGCTCTGCAAGGATCATCTCAATTTTGAAGAGTCTGCG 803
661 ATCGAAATCTAGCAGCTGAGACATCCACATATGAGAGAGAGCATTTTAAAGTGTG 720
804 ATCGAAATCTAGCAGCTGAGACATCCACATATGAGAGAGAGCATTTTAAAGTGTG 863
721 GTGAGAGAGAGAGGCTTCCGACATGCGCGCAAGTGAAGAGGCTCTGCAAGTTCAG 780
864 GTGAGAGAGAGAGGCTTCCGACATGCGCGCAAGTGAAGAGGCTCTGCAAGTTCAG 923
781 ATCAATGACTCGGTCACTCAATGAGAGCCCTTTTCTGTGAGAGACCTTATTTACTG 840
924 ATCAATGACTCGGTCACTCAATGAGAGCCCTTTTCTGTGAGAGATCTTATTTACTG 983
841 TCTCCAGCGGCTATCTTAAAGAGTGGAGAGAGGAGCTCAGGCTCAGCAGTGAAC 900
984 TCTCCAGCGGCTATCTTAAAGAGTGGAGAGAGGAGCTCAGGCTCAGCAGTGAAC 1043

QY 901 GATGGCCTCTCTTTTATCTCCAGTTCTGTATCATACCAACCAACACTGTTTGAACGT 960
DB 1044 GATGGCCTCTCTTTTATCTCCAGTTCTGTATCATACCAACCAACACTGTTTGAACGT 1103
QY 961 TCCATCTGGCCATCGCCGCGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
DB 1104 TCCATCTGGCCATCGCCGCGCTGATCTGTCTGCTGCTGCTGCTGCTGCTGCTG 1163
QY 1021 TGGTTTGGCCCTCTGCTGCTGCTGATGATTAAGAGAGTCCCTCCAGCCCTGCGAG 1080
DB 1164 TGGTTTGGCCCTCTGCTGCTGCTGATGATTAAGAGAGTCCCTCCAGCCCTGCGAG 1223
QY 1081 GAGAGTGAAGAAA 1094
DB 1224 GAGAGTGAAGAAA 1237

RESULT 7
AX458372 2112 bp DNA linear PAT 08-JUL-2002
LOCUS
DEFINITION
AX458372
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 Young, J.A., Bradley, K.A., Collier, R.J. and Mogridge, J.S.
AUTHORS
TITLE
JOURNAL
Patent: WO 0246228-A 7 13-JUN-2002;
WISCONSIN ALUMNI RESEARCH FOUNDATION (US)

FEATURES
source
1..2112
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
113..1114
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD38351.1"
/db_xref="GI:21725043"
/translation="MATARRALGIGFOWLSLATLVLICAGGGREBDGPGYGF
LYFLIDRSGVLLHMEIYFVEQLAKFISPOLRMSFIVFSTGTMKLTEDRGI
ROGLELOKYLPGGDTYMHGFERASQIYVENOGYRTASVIALTDGELHEDLFY
SERRANSRLGALIVCGVNDENRGLARIADSKHVPVNDGFOALOGIHSILKK
SCFIILAEBSSTICAGSFQVYVRGNFRRARAVDRVLCSPKINDSVTLNEKPSVSD
TYLCPAPILKEVGMKALDVSMDGLSPFISSVITITTHCSLHKIASGPTTAAME"

CDS

ORIGIN

Query Match 86.0%; Score 949.4; DB 6; Length 2112;
Best Local Similarity 99.9%; Pred. No. 2,66-272;
Matches 950; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGGCCAGCGGAGAGAGCCCTCGGCATCGGCTTCCAGTGGCTCTCTTTGGCCACT 60
DB 113 ATGGCCAGCGGAGAGAGCCCTCGGCATCGGCTTCCAGTGGCTCTCTTTGGCCACT 172
61 CTGGTCTCATCTGCGCGGCGGAGAGAGGAGATGGGGTTCAGCTGTAC 120
DB 173 CTGGTCTCATCTGCGCGGCGGAGAGAGGAGATGGGGTTCAGCTGTAC 232
121 GCGGATTTGACCTGTACTTATTTTGAAGAAATCAGAGAGTGTGACCACTGGAGAT 180
DB 233 GCGGATTTGACCTGTACTTATTTTGAAGAAATCAGAGAGTGTGACCACTGGAGAT 292
181 GAAATCTATTAATTGTGGAGAGATGGCTCAAAATTCATGAGCCCAAGTTGAGAAATG 240
DB 293 GAAATCTATTAATTGTGGAGAGATGGCTCAAAATTCATGAGCCCAAGTTGAGAAATG 352
QY 241 TCCTTTATTTGTTTCTCCACCCGAGAGAACCTTTAATGAACGTGACAGAGAGAGAA 300

Db	353	TCCTTATGTTTCTTCCACCCGAGAACACCTTAAAGAACTCAAGAAACAGAGAA	412
Qy	301	CAATCCCTCAAGGCTTGAAGAACTCCGAAAGTTCTGCGAGAGAGACACTACATG	360
Db	413	CAATCCCTCAAGGCTTGAAGAACTCCGAAAGTTCTGCGAGAGAGACACTACATG	472
Qy	361	CATGAGGATTTGAAAGGCGCAGTAGAGATTTATTAAGAAACAGCAAGGATACAG	420
Db	473	CATGAGGATTTGAAAGGCGCAGTAGAGATTTATTAAGAAACAGCAAGGATACAG	532
Qy	421	ACAGCAGGCTCATCTATGCTTCTTGACTGAGAGAACTCCATGAAGATCTCTTTTCTAT	480
Db	533	ACAGCAGGCTCATCTATGCTTCTTGACTGAGAGAACTCCATGAAGATCTCTTTTCTAT	592
Qy	481	TCAGAGAGGAGGCTTAAATAGGCTGAGATCTTGGAGCAATGTTACTGTTGATGAG	540
Db	553	TCAGAGAGGAGGCTTAAATAGGCTGAGATCTTGGAGCAATGTTACTGTTGATGAG	652
Qy	541	AAAGATTTCAATGAGACACAGCTGCGCCGAGATTGCGGACAGTAAAGATCATGTTTCC	600
Db	653	AAAGATTTCAATGAGACACAGCTGCGCCGAGATTGCGGACAGTAAAGATCATGTTTCC	712
Qy	601	GTGATGACGGCTTTCAGGCTTTCGAGGCTCATCTCAATTTTGAAGATCTCTGC	660
Db	713	GTGATGACGGCTTTCAGGCTTTCGAGGCTCATCTCAATTTTGAAGATCTCTGC	772
Qy	661	ATCGAATTTCTAGCAGCTGAACCATCCACATATGAGAGAGAGTCAATTCAGATTGTC	720
Db	773	ATCGAATTTCTAGCAGCTGAACCATCCACATATGAGAGAGAGTCAATTCAGATTGTC	832
Qy	721	GTGAGAGGAAACGGCTTTCGAGATGCGCGCAACGTGACAGGCTCTCTGACGTTCAAG	780
Db	833	GTGAGAGGAAACGGCTTTCGAGATGCGCGCAACGTGACAGGCTCTCTGACGTTCAAG	892
Qy	781	ATCATATGCTGGGTCACTCAATGAGAAAGCCCTTTCTGTGGAAGACACTTATTACTG	840
Db	893	ATCATATGCTGGGTCACTCAATGAGAAAGCCCTTTCTGTGGAAGACACTTATTACTG	952
Qy	841	TGTCCAGGCGCTATCTTAAAGAGTTGSCATGAAAGCTGCACTCAGGTACAGCATGAC	900
Db	953	TGTCCAGGCGCTATCTTAAAGAGTTGSCATGAAAGCTGCACTCAGGTACAGCATGAC	1012
Qy	901	GATGGCCTCTCTTTATCTCCAGTTCTGTCACTACACACACACTGT	951
Db	1013	GATGGCCTCTCTTTATCTCCAGTTCTGTCACTACACACACACTGT	1063
RESULT 8	2112 bp mRNA linear PRI 29-JUN-2004		
LOCUS	BC012074 Homo sapiens anhrax toxin receptor 1, transcript variant 3, mRNA		
DEFINITION	Homo sapiens MGC:15967 IMAGE:4563020, complete cde.		
ACCESSION	BC012074		
VERSION	BC012074.1 GI:15082332		
KEYWORDS	MGC.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens (human)		
REFERENCE	Eukaryota Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 2112)		
	Klausner R.D., Collins F.S., Wagner L.H., Shenmen G.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stepleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Udelin T.B., Toshitoki S., Carinici P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hult J.S.W., Villetton D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		

TITLE	Pahay, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywicki, M.I., Skalek, U., Smalls, D.E., Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)		
PUBMED	1247932		
REFERENCE	2 (bases 1 to 2112)		
AUTHORS	Strauberg, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (02-AUG-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov		
COMMENT	Contact: MGC help desk Email: cgabs-remail.nih.gov Tissue Procurement: DCTD/DRP cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@cgsc.bc.ca Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butlerfield, Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palamquist, Anca Petrescu, Anna Lisa Prabhu, Parvaneh Saeedi, Jr Santos, Angelique Scherch, Ursula Skalek, Duane Smalley, Jeff Scott, Miranda Tsai, George Yang, Jacque Schein, Asim Siddiqui, Rob Holt, Marco Marra.		
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAL Plate: 29 Row: 9 Column: 3 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gl: 16933552. Location/Qualifiers 1. .2112 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:19967 IMAGE:4563020" /issue_type="Kidney, renal cell adenocarcinoma" /clone_id="NIH_MGC_14" /lab_host="DH10B-R" /note="Vector: pOTB7" 1. .2112 /gene="ANTXR1" /note="Synonyms: FLJ10601, FLJ21776, TEN8, ATR" /db_xref="LocusID:84168" 113. .1114 /db_xref="MIM:606410" /gene="ANTXR1" /codon_start=1 /product="tumor endothelial marker 8, isoform 3 precursor" /protein_id="AAH12074.1" /db_xref="GI:15082333" /db_xref="MIM:606410" /translation="MATERALGIGROWSLATVVICAGGGRREDGPGACGGFD LFFLDSSGVLAHWNNTYIFVEQLAKFISPOURSFIVSTRGTTLMKLTBEROI ROGLIEQKVLPGDITMHEGFERASQIYENRQGRYSVIALTDGELHEDLFY SERRAANSRDLGAIIVCVGVDPIETQLARIADSKHVPFVNDGFMALOGIHSILK SCEIIAABSTICAGSPFVVVNGNFRARVNDVLCSEKINDSVTLNEKPSVSD TYLCPAPILKEVGMKALQVSNMDGSISSVITTTCSLHKIASGPTTAAQMG"		
gene	gene		
CDS	CDS		
ORIGIN	Query Match 86.0%; Score 949.4; DB 9; Length 2112; Best local similarity 99.9%; Pred. No. 2.6e-272;		

QY 804 TGAGAAAGCCCTTTCTGTGGAGAGACCTTATTACTGTGTCCAGCGCCCTATCTTAAAGA 863
 DB 1071 TGAGAAAGCCCTTTCTGTGGAGAGACCTTATTACTGTGTCCAGCGCCCTATCTTAAAGA 1130
 QY 864 AGTTGGCATGGAAGCTGCACTCCAGTGCAGTGAACGATGGCCCTCTCTTTATCTCCAG 923
 DB 1131 AGTTGGCATGGAAGCTGCACTCCAGTGCAGTGAACGATGGCCCTCTCTTTATCTCCAG 1190
 QY 924 TTCTGTCAATCATCAACCAACACACTGTGTCTGACGGTTCATCTGGCCATCGCCCTGCT 983
 DB 1191 TTCTGTCAATCATCAACCAACACACTGTGTCTGACGGTTCATCTGGCCATCTGCTGCT 1250
 QY 984 GATCTCTTCTGTCTCTTCAAGCCCTGCGCTCTCTGTGGTGTCTGCGCCCTCTGCTGAC 1043
 DB 1251 GGTCTCTTCTGTCTCTTCAAGCCCTGCGCTCTCTGTGGTGTCTGCGCCCTCTGCTGAC 1310
 QY 1044 TGTGATTCATCAAGAGAGTCCCTCCACCCCTGCGAGAGAGATGAGAAA 1094
 DB 1311 AGTATTCATCAAGAGAGTCCCTCCACCCCTGCGAGAGAGATGAGAAA 1361
 RESULT 10
 AX393370
 LOCUS AX393370 5220 bp DNA linear PAT 23-MAR-2002
 DEFINITION Sequence 300 from Patent WO0210217.
 ACCESSION AX393370
 VERSION AX393370.1 GI:19701356
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1
 AUTHORS St Croix B., Kinzler K.W. and Vogelstein B.
 TITLE Endothelial cell expression patterns
 JOURNAL Patent: WO 0210217-A 300 07-FEB-2002;
 The Johns Hopkins University (US)
 FEATURES
 source
 1..5220
 /organism="Mus musculus"
 /mol_type="unassigned DNA"
 /db_xref="taxon:10090"
 ORIGIN
 Query Match 77.3% Score 853.4; DB 6; Length 5220;
 Best Local Similarity 87.3%; Pred. No. 1.8e-243;
 Matches 935; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

DB 651 TGAGCAGATTTACTATGAGAACAGTCAAGATACAGACGCGAGCGTCATCATCGCGTT 710
 QY 444 GACGTATGAGAGACCTCCAGTAAGATCTCTTTTCAATCAGAGAGAGGCGCTAATAGTTC 503
 DB 711 GACGATGGAGAGCTGACAGAGACCTCTCTTCTTACTAGAGAGAGGCGCTAATAGTTC 770
 QY 504 TCGAGATCTGGTGCATTTGTTTACTGTGTGTGTGTGAAGATTCATATGAGACAGCT 563
 DB 771 CCGAGACCTTGGTGTGATTTGTTTACTGTGTGTGTGTGAAGATTCATATGAGACAGCT 830
 QY 564 GCGCCGATGCGGACAGTAAGATCATGTGTTCCTCGTAATGACGCGCTTACAGCTCT 623
 DB 831 GCGTCGATGTCAGACAGTAAGACAGAGTGTTCCTGTGAACGACGCGCTTACAGCTCT 890
 QY 624 GCAAGGATTCATCCACTCAATTTTGAAGAAGCTCGATCCGAATTCATAGACGTGAAC 683
 DB 891 CCAGGCAATTCATCCACTCAATTTTGAAGAAGCTCGATCCGAATTCATAGACGTGAAC 950
 QY 684 ATCCACCATATGTGACAGAGAGTCAATTCAGATTGTCTGAGAGAGAAACGCGCTTCCGACA 743
 DB 951 ATCCACCATATGTGACAGAGAGTCAATTCAGATTGTCTGAGAGAGAAATGCGTTCCGACA 1010
 QY 744 TGCCGCAACGTGACAGAGGCTCTCTGACGCTTCAAGATCAATGACTCGGTCACTGCA 803
 DB 1011 TGCCGCAATGTGACAGAGGCTCTCTGACGCTTCAAGATCAATGACTCGGTCACTGCA 1070
 QY 804 TGAGAAAGCCCTTTCTGTGGAGAGACCTTATTACTGTGTCCAGCGCCCTATCTTAAAGA 863
 DB 1071 TGAGAAAGCCCTTTCTGTGGAGAGACCTTATTACTGTGTCCAGCGCCCTATCTTAAAGA 1130
 QY 864 AGTTGGCATGGAAGCTGCACTCCAGTGCAGTGAACGATGGCCCTCTCTTTATCTCCAG 923
 DB 1131 AGTTGGCATGGAAGCTGCACTCCAGTGCAGTGAACGATGGCCCTCTCTTTATCTCCAG 1190
 QY 924 TTCTGTCAATCATCAACCAACACACTGTGTCTGACGGTTCATCTGGCCATCGCCCTGCT 983
 DB 1191 TTCTGTCAATCATCAACCAACACACTGTGTCTGACGGTTCATCTGGCCATCTGCTGCT 1250
 QY 984 GATCTCTTCTGTCTCTTCAAGCCCTGCGCTCTCTGTGGTGTCTGCGCCCTCTGCTGAC 1043
 DB 1251 GGTCTCTTCTGTCTCTTCAAGCCCTGCGCTCTCTGTGGTGTCTGCGCCCTCTGCTGAC 1310
 QY 1044 TGTGATTCATCAAGAGAGTCCCTCCACCCCTGCGAGAGAGATGAGAAA 1094
 DB 1311 AGTATTCATCAAGAGAGTCCCTCCACCCCTGCGAGAGAGATGAGAAA 1361
 RESULT 11
 AF378762
 LOCUS AF378762 5220 bp mRNA linear ROD 07-OCT-2001
 DEFINITION Mus musculus tumor endothelial marker 8 precursor (Tem8) mRNA,
 complete cds.
 ACCESSION AF378762
 VERSION AF378762.1 GI:15987504
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1
 AUTHORS Carson-Walter E.B., Vogelstein B., Kinzler K.W. and St. Croix B.
 TITLE Cell surface tumor endothelial markers are conserved in mice and
 humans
 JOURNAL Cancer Res. 61 (18), 6649-6655 (2001)
 MEDLINE 21443268
 PUBMED 11559528
 REFERENCE 2 (bases 1 to 5220)
 AUTHORS Carson-Walter E.B., Vogelstein B., Kinzler K.W. and St. Croix B.
 TITLE Direct Submission
 JOURNAL Submitted (10-MAY-2001) Oncology, Johns Hopkins University, 1650
 Orleans Street, Baltimore, MD 21231, USA

```
FEATURES
Source      Location/Qualifiers
gene        1..5220
            /organism="Mus musculus"
            /mol_type="mRNA"
            /db_xref="taxon:10090"
            1..5220
            /gene="Tem8"
            274..1962
            /gene="Tem8"
            /note="Tem8"
            /codon_start=1
            /product="tumor endothelial marker 8 precursor"
            /protein_id="AAL11999.1"
            /db_xref="GI:15987505"
            /translation="MDRAGRLGAGRLGCVAAALVLCAGHGREDGDPACYGFDLY
            FIDKSGSVLHNMNEIYFYVQLAHRTSPQIRMSFVFSRGTTLKLTEDREOIRQ
            GLEBQKVLPGSDTYMEGFERASQIYENSQRTASVIALTDGEHDLFPYSR
            REANRSRLGAIYVCVGVDPNEQLAIAISKDHVPEVNGFQALQIHSILKSKC
            IEILAAEPSTICAGSPQVVRNGNFRHARVNDVLCSEFKINDSVTLNEXFAVEDY
            LCPAPILKEVGMKALQVSMNDGLSFSSSVIITTHCSGSLIALALVFLILAL
            ALLMWFMLCTVIIKEVPPPVESSEEDDGLPKKMPVDAASYGSGVGIKRM
            EVRNGKSTSEGAKEKAKNARYMPQOEFPPEPRPLANNMRPSPRWKPIYG
            KLDALWILRKYDVSVMRPOPDGTCINFTRYKNSQPAKYPLUNNTYHSPPPAP
            IYTPPPAPHCPPAPSPAPPIBPSPSTLPPEPPAPPPNAPPSRPPPSV"
            274..354
            /gene="Tem8"
            355..1959
            /gene="Tem8"
            /product="tumor endothelial marker 8"
            /gene="Tem8"
            1228..1293
            /gene="Tem8"
            /note="Region: predicted transmembrane domain"
ORIGIN
Query Match      77.3%; Score 853.4; DB 10; Length 5220;
Best Local Similarity 87.3%; Pred. No. 1.8e-243;
Matches 935; Conservative 0; Mismatches 136; Indels 0; Gaps 0;
24 CCTGGCAATCGGCTTCAGTGGCTCTCTTTGGCACTCGTGGTCACTCTGCGCGGCA 83
Db |||||
QY |||||
291 CTGGGTGGGCGCTGCGGGACTTCGCTGCTGCACTCGTCTCTGTGGCGGCA 350
Db |||||
QY |||||
84 AGGGGAGCAGAGGAGATGGGGGTTCAGCTCTGCTACGCGGATTTGACTGTACTTCAT 143
Db |||||
QY |||||
351 CGGGGGCGCGCGAGATGGGGAACAAGCTTGTACGAGAGATTGCACTCTACTTCAT 410
Db |||||
QY |||||
144 TTTCGACAAATCAGAAATGTCTGCAACACTGGAATGAATCTTACTTTGTGGAACA 203
Db |||||
QY |||||
411 CCTGACAAGTCAGAAATGTCTGCAACACTGGAATGAATCTTACTTTGTGGAACA 470
Db |||||
QY |||||
204 GTTGGCTCAAAATTCATGAGCCCAAGTGAAGATGTCCTTATTTGTTTCTCCACCG 263
Db |||||
QY |||||
471 GTTGGCTCAAAATTCATGAGCCCAAGTGAAGATGTCCTTATTTGTTTCTCCACCG 530
Db |||||
QY |||||
264 AGGAACAACCTTATGAACCTGACAGAAAGAGAAACAATCCGTCAAGGCTTGAAGA 323
Db |||||
QY |||||
531 AGGACAACCTTATGAACCTGACAGAAAGAGAAACAATCCGTCAAGGCTTGAAGA 590
Db |||||
QY |||||
324 ACTCCAGAAAGTTCTGCCAGAGAGACCTTAATCATGCTGAAGATTTGAAGAGCCAG 383
Db |||||
QY |||||
591 GCTCCAGAAAGTTCTGCCAGAGAGACCTTAATCATGCTGAAGATTTGAAGAGCCAG 650
Db |||||
QY |||||
384 TGAGCAGATTTATTTATGAACAGACAAAGGTACAGAGCAGCGCATGATGCTT 443
Db |||||
QY |||||
651 TGAGCAGATTTATTTATGAACAGACAAAGGTACAGAGCAGCGCATGATGCTT 710
Db |||||
QY |||||
444 GACTGATGAGAACTTCATGAAGATCTTTTCTTATTTCAATGAGAGGCTTAATAGTC 503
Db |||||
QY |||||
711 GACGATGAGAGCTTCAGAGAACTTTCTTTCTTCACTCAAGAGAGGCTTAACGATC 770
Db |||||
QY |||||
504 TGAAGATCTTGGTCAATTTTATCTGTGTGTGTGAAGATTTCAATGAGACAGCT 563
Db |||||
QY |||||
771 CGAGACCTTGGTCAATTTTATCTGTGTGTGTGAAGATTTCAATGAGAACTCACTT 830
Db |||||
QY |||||
```

```
QY 564 GGCCTGGATTCGAGACAGTAAGATCATGTGTTCCGTAATGACGGCTTTCAGGCTCT 623
Db 831 GGCTCGGATTCGAGACAGTAAGACAGTGTGTTCCGTAATGACGGCTTTCAGGCTCT 890
QY 624 GCMAAGCATCATCCATCTTATTTGAAGAGTCTGATGGAATTTTACAGCTGAACC 683
Db 891 CCMAAGCATTCATCCATCTTATTTGAAGAGTCTGATGGAATTTTACAGCTGAACC 950
QY 684 ATCCACCATATGTCAGAGAGTCAATTTCAAGTTGTGTGAGAGAAACGGCTTCCGACA 743
Db 951 ATCCACCATCTGCGCGGAGAGTCTTCAAGTGTGTGTGAGAGAAATGGCTTCCGACA 1010
QY 744 TGCCCGCAAGTGAACAGAGTCTCTGCAAGCTTCAAGATCAATGATCTGCTCACTCAA 803
Db 1011 TGCCCGCAAGTGAACAGAGTCTCTGCAAGCTTCAAGATCAATGATCTGCTCACTCAA 1070
QY 804 TGAGAACGCTTTTCTGTGGAAGACATTTATTTACTGTGTCCAGCGCTTATTTAAAGA 863
Db 1071 TGAGAACGCTTTGCTGTGGAAGACATTTATTTGCTGTGCCAGACCAATCTTGAAGA 1130
QY 864 AGTTGGATGAAGCTGCACTCCAGTCAAGATGAACGATGAGCTCTCTTTATCTCCAG 923
Db 1131 AGTTGGATGAAGCTGCACTCCAGTCAAGATGAACGATGAGCTCTCTTTATCTCCAG 1190
QY 924 TTCTGTATCATCAACCAACAACAAGTGTGAAGGTTTCAATCTGAGCATGCGCTTCT 983
Db 1191 TTCTGTATCATCAACCAACAACAAGTGTGAAGGCTTCAATCTGAGCATGCGCTTCT 1250
QY 984 GATCTGTGCTGCTCTTCTAGCGCTGCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1043
Db 1251 GTCTCTTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1310
QY 1044 TGTGATATCAAGAGAGTCTCCACCCCTGCCAGAGAGAGTGAAGAAA 1094
Db 1311 AGTATATCAAGAGAGTCTCCACCCCTGCCAGAGAGAGTGAAGAAA 1361
RESULT 12
AR338517 1609 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 8 from patent US 6569662.
ACCESSION AR338517
VERSION AR338517.1 GI:33725374
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 1609)
Tang,Y.T., Zhou,P. and Drmanac,R.T.
Nucleic acids and polypeptides
JOURNAL
Patent: US 6569662-A 8 27-MAY-2003;
FEATURES
Source      Location/Qualifiers
gene        1..1609
            /organism="unknown"
            /mol_type="genomic DNA"
ORIGIN
Query Match      72.8%; Score 803.8; DB 6; Length 1609;
Best Local Similarity 99.8%; Pred. No. 9.8e-229;
Matches 805; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
1 ATGGCAGCGGCGAGAGAGCCCTCGGCAATCGGCTTCCAGTGGCTCTCTTTGGCACT 60
Db |||||
QY |||||
309 ATGGCAGCGGCGAGAGAGCCCTCGGCAATCGGCTTCCAGTGGCTCTCTTTGGCACT 368
Db |||||
QY |||||
61 CTGGTCTCATCTGGCGCGGCAAGGGAGCGCAGGGAGATGGGGGTTCAGGCTGTAC 120
Db |||||
QY |||||
369 CTGGTCTCATCTGGCGCGGCAAGGGAGCGCAGGGAGATGGGGGTTCAGGCTGTAC 428
Db |||||
QY |||||
121 GCGGATTTGACCTGATCTTCAATTTTGAACAATCAGGAAGTGTGCTGACCACTGGAAT 180
Db |||||
QY |||||
429 GCGGATTTGACCTGATCTTCAATTTTGAACAATCAGGAAGTGTGCTGACCACTGGAAT 488
Db |||||
QY |||||
```

QY 181 GAAATCTATTATTGTGGACAGTGGCTCAAAATTCATGACCCACAGTTGGAATG 240
DB 489 GAAATCTATTATTGTGGACAGTGGCTCAAAATTCATGACCCACAGTTGGAATG 548
QY 241 TCTTTATTGTTTCTCCACCCGAGAGACACTTATGAACTGACAGAGACAGAGA 300
DB 549 TCTTTATTGTTTCTCCACCCGAGAGACACTTATGAACTGACAGAGACAGAGA 608
QY 301 CAATCCGTCAGAGGCTAGAGAACTCCAGAAAGTTCCGAGAGAGACACTTACATG 360
DB 609 CAATCCGTCAGAGGCTAGAGAACTCCAGAAAGTTCCGAGAGAGACACTTACATG 668
QY 361 CATGAGAGATTGAAAGGCGCAGTGCAGATTTATATGAAACAGACAAAGGTACAG 420
DB 669 CATGAGAGATTGAAAGGCGCAGTGCAGATTTATATGAAACAGACAAAGGTACAG 728
QY 421 ACAGCCAGGCTCATCTGCTTTGACTGATGAGAACTCCATGAAGATCTTTTCTAT 480
DB 729 ACAGCCAGGCTCATCTGCTTTGACTGATGAGAACTCCATGAAGATCTTTTCTAT 788
QY 481 TCAGAGAGGAGGCTAATAGTCTGAGATCTTGAGCAATGTTTACTGTGTGTG 540
DB 789 TCAGAGAGGAGGCTAATAGTCTGAGATCTTGAGCAATGTTTACTGTGTGTG 848
QY 541 AAAGATTTCAATGAGACACAGCTGGCCCGGATTTGGGACAGTAAAGATCTGTTCC 600
DB 849 AAAGATTTCAATGAGACACAGCTGGCCCGGATTTGGGACAGTAAAGATCTGTTCC 908
QY 601 GTGAATGACGGCTTTCAGGCTCTGCAAGGATCATCTCAATTTTGAAGAGTCTGC 660
DB 909 GTGAATGACGGCTTTCAGGCTCTGCAAGGATCATCTCAATTTTGAAGAGTCTGC 968
QY 661 ATCGAAATTTAGAGAGCTGAACCAATATGAGAGAGATCTTCAAGTGTG 720
DB 969 ATCGAAATTTAGAGAGCTGAACCAATATGAGAGAGATCTTCAAGTGTG 1028
QY 721 GTGAGAGGAAACGGCTTCCGACATGCCGCAACGTGACAGGGTCTCTGACGTTCAAG 780
DB 1029 GTGAGAGGAAACGGCTTCCGACATGCCGCAACGTGACAGGGTCTCTGACGTTCAAG 1088
QY 781 ATCAATGACTCGGTCACTCACTCAATGAG 807
DB 1089 ATCAATGACTCGGTCACTCACTCAATGAG 1115

RESULT 13
LOCUS CQ498433 1718 bp DNA linear PAT 30-JAN-2004
DEFINITION Sequence 30300 from Patent WO0160860.
ACCESSION CQ498433
VERSION CQ498433.1 GI:41464069
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
Schlegel R., Endege W.O. and Monahan J.E.
Genes differentially expressed in human prostate cancer and their
use
Patent: WO 0160860-A 30300 23-AUG-2001;
JOURNAL Millennium Predictive Medicine, Inc. (US)
FEATURES
LOCATION/Qualifiers
SOURCE 1..1718
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 72.7%; Score 802.2; DB 6; Length 1718;
Best Local Similarity 99.6%; Pred. No. 3e-228;
Matches 804; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGCCAGGCGAGAGCGAGAGCCCTCGGCATCGGCTTCAAGTGGCTCTTTGGCCACT 60
DB 413 ATGGCCAGGCGAGAGCGAGAGCCCTCGGCATCGGCTTCAAGTGGCTCTTTGGCCACT 472
QY 61 CTGGTCTCATCTGCGCGCGGAGAGGAGAGAGAGATGGGGTTCAGCCTGTAC 120
DB 473 CTGGTCTCATCTGCGCGCGGAGAGGAGAGAGAGATGGGGTTCAGCCTGTAC 532
QY 121 GCGGATTTGACCTGTACTTCAATTTTGAACAATCAGAAATGTGTCTCACTGAGAT 180
DB 533 GCGGATTTGACCTGTACTTCAATTTTGAACAATCAGAAATGTGTCTCACTGAGAT 592
QY 181 GAAATCTATTATTGTGGAACAGTGGCTCAAAATTCATCAGCCACAGTTGGAATG 240
DB 593 GAAATCTATTATTGTGGAACAGTGGCTCAAAATTCATCAGCCACAGTTGGAATG 652
QY 241 TCTTTATTGTTTCTCCACCCGAGAGACACTTATGAACTGACAGAAACAGAGA 300
DB 653 TCTTTATTGTTTCTCCACCCGAGAGACACTTATGAACTGACAGAAACAGAGA 712
QY 301 CAATCCGTCAGAGGCTAGAGAACTCCAGAAAGTTCCGAGAGAGACACTTACATG 360
DB 713 CAATCCGTCAGAGGCTAGAGAACTCCAGAAAGTTCCGAGAGAGACACTTACATG 772
QY 361 CATGAGAGATTGAAAGGCGCAGTGCAGATTTATATGAAACAGACAAAGGTACAG 420
DB 773 CATGAGAGATTGAAAGGCGCAGTGCAGATTTATATGAAACAGACAAAGGTACAG 832
QY 421 ACAGCCAGGCTCATCTGCTTTGACTGATGAGAACTCCATGAAGATCTTTTCTAT 480
DB 833 ACAGCCAGGCTCATCTGCTTTGACTGATGAGAACTCCATGAAGATCTTTTCTAT 892
QY 481 TCAGAGAGGAGGCTAATAGTCTGAGATCTTGAGCAATGTTTACTGTGTGTG 540
DB 893 TCAGAGAGGAGGCTAATAGTCTGAGATCTTGAGCAATGTTTACTGTGTGTG 952
QY 541 AAAGATTTCAATGAGACACAGCTGGCCCGGATTTGGGACAGTAAAGATCTGTTCC 600
DB 953 AAAGATTTCAATGAGACACAGCTGGCCCGGATTTGGGACAGTAAAGATCTGTTCC 1012
QY 601 GTGAATGACGGCTTTCAGGCTCTGCAAGGATCATCTCAATTTTGAAGAGTCTGC 660
DB 1013 GTGAATGACGGCTTTCAGGCTCTGCAAGGATCATCTCAATTTTGAAGAGTCTGC 1072
QY 661 ATCGAAATTTAGAGAGCTGAACCAATATGAGAGAGATCTTCAAGTGTG 720
DB 1073 ATCGAAATTTAGAGAGCTGAACCAATATGAGAGAGATCTTCAAGTGTG 1132
QY 721 GTGAGAGGAAACGGCTTCCGACATGCCGCAACGTGACAGGGTCTCTGACGTTCAAG 780
DB 1133 GTGAGAGGAAACGGCTTCCGACATGCCGCAACGTGACAGGGTCTCTGACGTTCAAG 1192
QY 781 ATCAATGACTCGGTCACTCACTCAATGAG 807
DB 1193 ATCAATGACTCGGTCACTCACTCAATGAG 1219

RESULT 14
LOCUS BD156323 1436 bp DNA linear PAT 17-JAN-2003
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD156323
VERSION BD156323.1 GI:27862081
KEYWORDS JP 2002191363-A/11166.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1436)
Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
Primer for synthesizing full-length cDNA and use thereof

ORIGIN	FEATURES	source
Query Match	71.3%; Score 787; DB 6; Length 1436;	
Best Local Similarity	99.3%; Pred. No. 1e-223;	
Matches	801; Conservative 0; Mismatches 5; Indels 1; Gaps 1.	
Db	1 ATGGCCACGGCGGAGCGGAGAGACCTCTCGGCATCGGCTTCATAGTGGCTCTCTTTGGCCACT 60	
Db	144 ATGGCCACGGCGGAGCGGAGAGACCTCTCGGCATCGGCTTCATAGTGGCTCTC-ACGGCCACT 202	
Qy	61 CTGTGCTCATCTGGCGCGGCAAGGAGGAGCCGACGAGAGATGGGAGTTCACGCTGTAC 120	
Db	203 CTGTGCTCATCTGGCGCGGCAAGGAGGAGCCGACGAGAGATGGGAGTTCACGCTGTAC 262	
Qy	121 GCGGAGTTTGACCTGTACTTCATTTTGGACAATATCAGAAAGTGTGTGTGACCATGTGAAT 180	
Db	263 GCGGAGTTTGACCTGTACTTCATTTTGGACAATATCAGAAAGTGTGTGTGACCATGTGAAT 322	
Qy	181 GAAATCTATTTACTTTGTGGAAAGATGGGTCAAAATCATAGGCCACAGTTGAATG 240	
Db	323 GAAATCTATTTACTTTGTGGAAAGATGGGTCTCAAAATTCATACGCCACAGTTGAATG 382	
Qy	241 TCTTTATTTGTTTCTTCACCCGAGAGACAACTTAATAAACTGACAGAGACAGAGAA 300	
Db	383 TCTTTATTTGTTTCTTCACCCGAGAGACAACTTAATAAACTGACAGAGACAGAGAA 442	
Qy	301 CAAATCCGTCAAGGCTTGAAGAACTTCAGAAAGTTCTGCCAGAGAGACACTTAATG 360	
Db	443 CAAATCCGTCAAGGCTTGAAGAACTTCAGAAAGTTCTGCCAGAGAGACACTTAATG 502	
Qy	361 CATGAAGATTTGAAAGGCGCAGTGGAGAGATTTATATGAAACAGACAAAGGATACAG 420	
Db	503 CATGAAGATTTGAAAGGCGCAGTGGAGAGATTTATATGAAACAGACAAAGGATACAG 562	
Qy	421 ACAGCCAGGTCATCTTGTGAAGTGAAGAACTTCATGAAGATCTCTTTTCTAT 480	
Db	563 ACAGCCAGGTCATCTTGTGAAGTGAAGAACTTCATGAAGATCTCTTTTCTAT 622	
Qy	481 TCAGAGAGGAGGCTAATAGTCTGAGATCTTTGGGCAATTTGTTACGTGTGTGTG 540	
Db	623 TCAGAGAGGAGGCTAATAGTCTGAGATCTTTGGGCAATTTGTTACGTGTGTGTG 682	
Qy	541 AAAAGATTTCAATGAGACAGAGTGGCCCGGATTTGCGACAGTAAGGATCATGTGTTCC 600	
Db	683 AAAAGATTTCAATGAGACAGAGTGGCCCGGATTTGCGACAGTAAGGATCATGTGTTCC 742	
Qy	601 GTGAATGACGGCTTTAGGCTCTGCAAGGATCATCACTCAATTTTGAAGAGTCTG 660	
Db	743 GTGAATGACGGCTTTAGGCTCTGCAAGGATCATCACTCAATTTTGAAGAGTCTG 802	
Qy	661 ATCGAAATCTGAGAGTCAACCATCACTATGTGACGAGAGAGTCAATTAAGTGTG 720	

Db		803	ATCGAAATCTTAGCAGGTGGAACCATTCACCACTAATGTGCAGAGAAGTCAATTCAAAGTTGTC	862
Oy		721	GTCAGAGAAAACGGCTTCCGACATATGCCCGCAAGCTGGAACAGGGTCTCTTGACGTTCAAG	780
Db		863	GTCAGAGAAAACGGCTTCCGACATATGCCCGCAAGCTGGAACAGGGTCTCTTGACGTTCAAG	922
Oy		761	ATCATGACTCGGTCACTCACTCAATGAG	807
Db		923	ATCATGACTCGGTCACTCACTCAATGAG	949
RESULT_15				
LOCUS	AX458374	1436 bp	DNA	linear PAT 08-JUN-2002
DEFINITION	Sequence 9 from Patent WO0246228.			
ACCESSION	AX458374			
VERSION	AX458374.1			
KEYWORDS	GI:21725044			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euleleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 Young,J.A., Bradley,K.A., Collier,R.J. and Mogridge,J.S. Receptor for b. Anthracis toxin Patent: WO 0246228-A 9 13-JUN-2002;			
AUTHORS	MISCONSIN ALUMNI RESEARCH FOUNDATION (US)			
TITLE	Location/Qualifiers			
JOURNAL	1..1436 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606" 380..1036 /note="unnamed protein product" /codon_start=1 /protein_id="CAD38352.1" /db_xref="GI:21725045" /translation="MSFIVSTRTGLTKLTEDREQIRQLGLELQKVLPGSDTYMHREG PERASEOIYVENRGVRYASVILALTDGELEHEDLPFYSERBANSRDLAGIIVGVGW DFNETOLARIADSKDHVPVDVFOALOGIHSILKSCEIIILAESTTCAGESFPVW VVRNGSPRPHANVDNRVLCFSFRINDSVLSLSLGPMWSSTSGFGKNSHPCLPARPHT			
CDS				
ORIGIN				
Query Match	71.3%; Score 787; DB 6; Length 1436;			
Best Local Similarity	99.3%; Pred. No. 1e-223;			
Matches	801; Conservative 0; Mismatches 5; Indels 1; Gaps 1;			
Oy		1	ATGGCCACGGGGAGCGGAGAGCCCTCGGACATCGGCTTCCAGTGGCTCTTTGGCCACT	60
Db		144	ATGGCCACGGGGAGCGGAGAGCCCTCGGACATCGGCTTCCAGTGGCTCTTTGGCCACT	202
Oy		61	CTGTGTCATCTCTGCGCGGGCAAGGGGGAAGCAGAGGAGATGGGGGCTCACGCTGTAC	120
Db		203	CTGTGTCATCTCTGCGCGGGCAAGGGGGAAGCAGAGGAGATGGGGGCTCACGCTGTAC	262
Oy		121	GGCGGATTGACCTGTACTTCATTTTGGACAATAATCAGGAAGTGTGCTGCACCACTGAAT	180
Db		263	GGCGGATTGACCTGTACTTCATTTTGGACAATAATCAGGAAGTGTGCTGCACCACTGAAT	322
Oy		181	GAAATCTTTACTTTGTGAGAACAGTGGCTCAAAATTCATCAGCCCCAGTTGAGATG	240
Db		323	GAAATCTTTACTTTGTGAGAACAGTGGCTCAAAATTCATCAGCCCCAGTTGAGATG	382
Oy		241	TCCTTTAATGTTTCTCCACCAGGAACAACCTTAATGAACCTGACAGAAAGACAGAA	300
Db		383	TCCTTTAATGTTTCTCCACCAGGAACAACCTTAATGAACCTGACAGAAAGACAGAA	442
Oy		301	CAAAATCCGCAAGGCTTGAAGAACTCCAGAAAGTTCTGCGACAGAGAGACATTTACATG	360
Db		443	CAAAATCCGCAAGGCTTGAAGAACTCCAGAAAGTTCTGCGACAGAGAGACATTTACATG	502

```
Oy 361 CATGAAGATTGAAAGGGCCAGTGAAGATTTATATGAAAACAGACAGGGTACAGG 420
    |||
Db 503 CATGAAGATTGAAAGGGCCAGTGAAGATTTATATGAAAACAGACAGGGTACAGG 562
    |||
Oy 421 ACAGCAGCGGTATCATTTGCTTTGACTGTGAGAACTCCATGAAGATCTTTTTCTAT 480
    |||
Db 563 ACAGCTACGCTCATTTGCTTTGACTGTGAGAACTCCATGAAGATCTTTTTCTAT 622
    |||
Oy 481 TCAGAGAGGGAGCTTAAATAGATCTCGAGATCTTGAGCAATTGTTACTGTGTGTG 540
    |||
Db 623 TCAGAGAGGGAGCTTAAATAGATCTTGAGATCTTGAGCAATTGTTACTGTGTGTG 682
    |||
Oy 541 AAAGATTTCAATGAGACACAGCTGGCCCGAATTCGGAAGATCATGTGTTCCC 600
    |||
Db 683 AAAGATTTCAATGAGACACAGCTGGCCCGAATTCGGAAGATCATGTGTTCCC 742
    |||
Oy 601 GTGAATGACGGCTTTCAGGCTCTGCAAGGCATCATCTCAATTTTGAAAGATCTGC 660
    |||
Db 743 GTGAATGACGGCTTTCAGGCTCTGCAAGGCATCATCTCAATTTTGAAAGATCTGC 802
    |||
Oy 661 ATCGAAATTTCTAGCAGCTGAACCATCCATATATGTGAGAGAGTCAATTCAGTTGTC 720
    |||
Db 803 ATCGAAATTTCTAGCAGCTGAACCATCCATATATGTGAGAGAGTCAATTCAGTTGTC 862
    |||
Oy 721 GTGAGAGGAAACGGCTTCCGACATGCCGCAACGTGACAGGGTCTCTGCAGCTTCAAG 780
    |||
Db 863 GTGAGAGGAAACGGCTTCCGACATGCCGCAACGTGACAGGGTCTCTGCAGCTTCAAG 922
    |||
Oy 781 ATCAATGACTCGGTCACTCAATGAG 807
    |||
Db 923 ATCAATGACTCGGTCACTCAATGAG 949
    |||
```

Search completed: June 14, 2005, 15:32:17
Job time : 5150 secs

THIS PAGE BLANK (USPTO)

xx The present sequence is that of cDNA encoding a human surface-bound anthrax toxin receptor (ATR). The cDNA is a PCR amplification product from HeLa cells and human placenta cDNA libraries. Anthrax toxin protective antigen (PA) binds to the ATR at a von Willibrand factor A domain located in the extracellular domain of ATR. The invention provides ATR polypeptides and polynucleotides, vectors, host cells, and transgenic and knock-out animals. The invention also provides methods for identifying molecules that bind the ATR and which reduce the toxicity of anthrax toxin. A claimed method for treating anthrax in a human or animal involves administering an agent that inhibits binding between PA and ATR at a level effective to reduce the severity of anthrax. Suitable agents include ATR or a PA-binding fragment of ATR, a PA-binding polypeptide at least 80% identical to these, a fusion protein, a monoclonal or polyclonal antibody, a polysaccharide, a lipid or a nucleic acid. ATR polynucleotides can also be used in the recombinant production of ATR polypeptides, and as molecular probes

SQ Sequence 1414 BP; 394 A; 344 C; 344 G; 332 T; 0 U; 0 Other;

Query Match 100.0%; Score 1104; DB 6; Length 1414;

```

      GC Content: 20.68; GC bias: no; 0;
      Matches 1104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY	ATGGCCACCGCCGAGCGAGAGCCCTCCGATCGGCTCCAGCGGCTCTTTGGCACT	6
Db	ATGGCCACCGCCGAGCGAGAGCCCTCCGATCGGCTCCAGCGGCTCTTTGGCACT	104
QY	CTGGTGCTCATCTGCGCCGGGGCAAGGGGGACGAGGAGGATGGGGGCTCAGCTCTAC	120
Db	CTGGTGCTCATCTGCGCCGGGGCAAGGGGGACGAGGAGGATGGGGGCTCAGCTCTAC	164
QY	GGCGGATTTGACCTGACTTCATTTTTGGACAAATCAGAAAGTGTGTGCACACTGGAAT	180
Db	GGCGGATTTGACCTGACTTCATTTTTGGACAAATCAGAAAGTGTGTGCACACTGGAAT	224
QY	GAAATCTATCTTTGTGGAAACAATTGTCGACCAATTCAGAAAGTGTGTGCACCACTGGAAT	240
Db	GAAATCTATCTTTGTGGAAACAATTGTCGACCAATTCAGAAAGTGTGTGCACCACTGGAAT	284
QY	TCCTTTATTTGTTTCTCCACCCGAGGACAACTTATGAAACTGACAGAAACAGAGAA	300
Db	TCCTTTATTTGTTTCTCCACCCGAGGACAACTTATGAAACTGACAGAAACAGAGAA	344
QY	CAAAATCCGTCAGAGGCTTAAGAAATCTCCAGAAAGTCTGCACAGAGACACTTACATG	360
Db	CAAAATCCGTCAGAGGCTTAAGAAATCTCCAGAAAGTCTGCACAGAGACACTTACATG	404
QY	CATGAAGATTTGAAAGGGCCAGTGTGACAGATTTATATGAAACAGACAAAGGCTACAG	420
Db	CATGAAGATTTGAAAGGGCCAGTGTGACAGATTTATATGAAACAGACAAAGGCTACAG	464
QY	ACAGCCAGGTCATCATCTTGTGACTGATGAGAACTCCATGAAGATCTCTTTTCTAT	480
Db	ACAGCCAGGTCATCATCTTGTGACTGATGAGAACTCCATGAAGATCTCTTTTCTAT	524
QY	TCAGAGAGGAGGCTAATAGTCTCGAAGATCTTGTGCAATTGTTTATCTGTGTGGTGG	540
Db	TCAGAGAGGAGGCTAATAGTCTCGAAGATCTTGTGCAATTGTTTATCTGTGTGGTGG	584
QY	AAAGATTTCAATGAGACACAGCTGGCCGGATTGCCAGACGTAAAGATCATGTGTTCCC	600
Db	AAAGATTTCAATGAGACACAGCTGGCCGGATTGCCAGACGTAAAGATCATGTGTTCCC	644
QY	GTTGAATGACGGCTTTCAAGGCTCTGCAAGGCATATCACTCAATTTTGAAGAAAGTCTGC	660
Db	GTTGAATGACGGCTTTCAAGGCTCTGCAAGGCATATCACTCAATTTTGAAGAAAGTCTGC	704
QY	ATCGAAATTTCTAGACACTGAAACCATCAACATATGTGACAGAGATCATTTCAAGTTGTC	720
Db	ATCGAAATTTCTAGACACTGAAACCATCAACATATGTGACAGAGATCATTTCAAGTTGTC	764
QY	GTGAGAGAAACGGCTTTCCGACATGTCCCGCAACGTGGACAGGGTCTCTTGACACTTCAG	780

Db	824	GTGAGAGAAACGGCTTCCGACATGCCCCGCACTGGACAGGGTCTCTGTGAGTTTCAG	883
Qy	781	ATCAATGACTCGGTCACTCATGAGAGAGCCCTTTTCTGTGAAGACATTATTACTG	840
Db	884	ATCAATGACTCGGTCACTCAATGAGAGCCCTTTTCTGTGAAAGACATTATTACTG	943
Qy	841	TGTCCAGGCGCTATCTTAAAAAGAGTTGGCATGAAGCTGACATCCAGGTGAGATGAC	900
Db	944	TGTCCAGGCGCTATCTTAAAGAGTTGGCATGAAGCTGACATCCAGGTGAGATGAC	1003
Qy	901	GATGAGCCTCTCTTTATCTCCAGTTCTGTCAATCAACACACACACTGTTCGTGACGT	960
Db	1004	GATGAGCCTCTCTTTATCTCCAGTTCTGTCAATCAACACACACTGTTCGTGACGT	1063
Qy	961	TCGATCTCGGCGCATCGCCCTGCTGATCTGTTCCTGTCCAGCCCTGGCTCTCTCTCG	1020
Db	1064	TCGATCTCGGCGCATCGCCCTGCTGATCTGTTCCTGTCCAGCCCTGGCTCTCTCTCG	1123
Qy	1021	TGGTTCTGGCCCTCTGCTGCACTGTGATTATCAAGAGGTCCTTCACCCCTGCGGAG	1080
Db	1124	TGGTTCTGGCCCTCTGCTGCACTGTGATTATCAAGAGGTCCTTCACCCCTGCGGAG	1183
Qy	1081	GAGAGTGAAGAAATTAATTAATAA	1104
Db	1184	GAGAGTGAAGAAATTAATTAATAA	1207

RESULT 2

ABL92078 standard; cDNA; 5540 BP.

AC ABL92078;

DT 30-MAY-2002 (first entry)

Human Tumour Endothelial Marker polynucleotide SEQ ID NO 176.

KM Human; mouse; rat; TEM; tumour endothelial marker; NEM; PEM; cytosstatic;
 KM normal endothelial marker; pan-endothelial marker; immunostimulant;
 KM antiangiogenic; tumour; neoangiogenesis; vascularised tumour;
 KM polycystic kidney disease; diabetes; retinopathy; rheumatoid arthritis;
 KM psoriasis; gene; ss.

Homo sapiens.

PN WO200210217-A2

PD 07-FEB-2002.

01-AUG-2001; 2001WO-US024031.

PR 02-AUG-2000; 2000US-0222599P.

PR 11-APR-2001; 2001US-0282850P.

PA (UYJO) UNIV JOHNS HOPKINS.

PI St Croix B, Kinzler KW, Vogelstein B;

DR WPI; 2002-291856/33.

PT An isolated molecule comprising an antibody variable region which

PT marker (TEM) protein, useful for inhibiting tumor growth.

PS Disclosure; Page 121-123; 331pp; English.

CC The invention relates to an isolated molecule comprising an antibody

CC tumour endothelial marker (TEM) protein selected from ABB90732, ABB90740,

CC proteins have cytostatic, immunostimulant and antiangiogenic activity.

OY	1	TTGGCGACGGGCGGAGCGGAGAGCCCTGGACATCGGCTTCCAGTGGCTCTCTTTGGCACT	60
Db	144	ATGGCGACGGGCGGAGCGGAGAGCCCTGGACATCGGCTTCCAGTGGCTCTCTTTGGCACT	203
OY	61	CTGGGCTCATCTGGCGCGGCGCAAGGGGGAAGCAGAGAGAGATGGGGGTCCAGCCTGTAC	120
Db	204	CTGGGCTCATCTGGCGCGGCGCAAGGGGGAAGCAGAGAGAGATGGGGGTCCAGCCTGTAC	263
OY	121	GGCGGATTTGACCTGTACTTCAATTTTGGACAAATCAGGAAGTGTCTGCACCATGGAAAT	180
Db	264	GGCGGATTTGACCTGTACTTCAATTTTGGACAAATCAGGAAGTGTCTGCACCATGGAAAT	323
OY	181	GAAATCTTATCTTTGTGGAAACAGTGGCTCAAAATTATCAGGCCACAGTTGGAATG	240
Db	324	GAAATCTTATCTTTGTGGAAACAGTGGCTCAAAATTCATCAGGCCACAGTTGGAATG	383
OY	241	TCCCTTATTTGTTTCTCCACCCGAGGAACAACCTTAATGAACTACAGAAACAGAGAA	300
Db	384	TCCCTTATTTGTTTCTCCACCCGAGGAACAACCTTAATGAACTACAGAAACAGAGAA	443
OY	301	CAAAATCCGTCAAGGCGCTAGAAAGAACTCCAGAAAGTTCTGCAGAGAGAGACATTACATG	360
Db	444	CAAAATCCGTCAAGGCGCTAGAAAGAACTCCAGAAAGTTCTGCAGAGAGAGACATTACATG	503
OY	361	CATGAAGATTTGAAAGGGGCCAGTGAAGATTTTATGAAACAGACAAAGGGTACAG	420
Db	504	CATGAAGATTTGAAAGGGGCCAGTGAAGATTTTATGAAACAGACAAAGGGTACAG	563
OY	421	ACAGCCAGCGTCAATCATTTGCTTGACTGATGGAGAACTCAATGAAGATCTCTTTTCTAT	480
Db	564	ACAGCCAGCGTCAATCATTTGCTTGACTGATGGAGAACTCAATGAAGATCTCTTTTCTAT	623
OY	481	TCAGAGAGGAGGCTAATAGTCTTCGAGATCTTGGTGCAAATGTTTACTGTGTGGTGG	540
Db	624	TCAGAGAGGAGGCTAATAGTCTTCGAGATCTTGGTGCAAATGTTTACTGTGTGGTGG	683
OY	541	AAAGATTTCAATGAGACACAGCTGGCCCGGATTTGGGACAGTAAGATCATGTGTTTCCC	600
Db	684	AAAGATTTCAATGAGACACAGCTGGCCCGGATTTGGGACAGTAAGATCATGTGTTTCCC	743
OY	601	GTGAATGACGGCTTTCAGGCTCTGGAGGACATCATCCACTCAATTTTGAAGAGTCTGC	660
Db	744	GTGAATGACGGCTTTCAGGCTCTGGAGGACATCATCCACTCAATTTTGAAGAGTCTGC	803
OY	661	ATCGAAATTTAGACGCTGAACCATCCAACATATGTGACGAGAGATCAATTCAGTTGTC	720
Db	804	ATCGAAATTTAGACGCTGAACCATCCAACATATGTGACGAGAGATCAATTCAGTTGTC	863
OY	721	GTGAGAGGAAACGGCTTCCGACATGCCCGCAACGTGACAGGGTCTCTGCAGCTTCAAG	780
Db	864	GTGAGAGGAAACGGCTTCCGACATGCCCGCAACGTGACAGGGTCTCTGCAGCTTCAAG	923
OY	781	ATCAATGACTGGGTCACTCAATGAGAGCCCTTTCTGTGGGAAGACCTAATTTATACG	840
Db	924	ATCAATGACTGGGTCACTCAATGAGAGCCCTTTCTGTGGGAAGATCTTAATTTACTG	983
OY	841	TGTCAGAGCGCTATCTTAAAGAAAGTTGGACATGAAGCTGCATCCAGGTCCAGATGAAC	900
Db	984	TGTCAGAGCGCTATCTTAAAGAAAGTTGGACATGAAGCTGCATCCAGGTCCAGATGAAC	1043
OY	901	GATGGCTCTCTTTTATCTCCAGTTCTGTCAATCATCAACACACACACTGTTCAGAGGT	960
Db	1044	GATGGCTCTCTTTTATCTCCAGTTCTGTCAATCATCAACACACACACTGTTCAGAGGT	1103
OY	961	TCCATCTCGGCAATGGCCCTGTGATCTCGTATCTCGTCTGAGCCCTGGCTCTCTCTCG	1020
Db	1104	TCCATCTCGGCAATGGCCCTGTGATCTCGTATCTCGTCTGAGCCCTGGCTCTCTCTCG	1163
OY	1021	TGGTCTGGCCCTCTGTGTCGACTGTGATTAATCAAGAGAGTCCCTCCACCCCTGCCAG	1080
Db	1164	TGGTCTGGCCCTCTGTGTCGACTGTGATTAATCAAGAGAGTCCCTCCACCCCTGCCAG	1222
OY	1081	GAGAGTGAAGAAAAA	1094

DB	1224	GAGAGTGAAGAGA	1237
RESULT 6			
ABX72003			
ID	ABX72003	standard; DNA; 5540 BP.	
AC	ABX72003;		
XX			
DT	12-MAR-2003	(first entry)	
XX			
DE	DNA encoding human tumour endothelial marker TEM 8.		
XX			
KM	Human; endothelial cell; EC; tumour endothelial cell; TEM; NEM;		
KM	Tumour endothelial marker; normal endothelial marker; PEM;		
KM	pan-endothelial marker; polycystic kidney disease; psoriasis;		
KM	diabetic retinopathy; rheumatoid arthritis; tumour angiogenesis;		
KM	neovascularization; immune response; cytoskeletal; gene;		
KM	ophthalmological; antineoplastic; antirheumatic; antipsoriatic; ds.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200283874-A2.		
XX			
PD	24-OCT-2002.		
XX			
PF	10-APR-2002; 2002WO-US008253.		
XX			
PR	11-APR-2001; 2001US-0282850P.		
PR	06-FEB-2002; 2002US-0354262P.		
XX			
PA	(UJJO) UNIV JOHNS HOPKINS.		
XX			
PI	Carson-Walter E, St Croix B, Kinzler KM, Vogelstein B;		
DR	WPI; 2003-093016/08.		
DR	P-PSDB; AB954430.		
PT	New purified human transmembrane protein, designated as tumor endothelial		
PT	marker (TEM) 3, useful for detecting, diagnosing or treating tumors,		
PT	polycystic kidney disease, diabetic retinopathy, rheumatoid arthritis or		
PS	psoriasis.		
XX			
PS	Disclosure: Page 117-120; 374pp; English.		
XX			
CC	The present invention relates to a novel method for the isolation of		
CC	endothelial cells (ECs), and the identification of genes expressed in		
CC	normal and tumour ECs. Tumour endothelial marker (TEM), normal		
CC	endothelial marker (NEM), and pan-endothelial marker (PEM) genes are		
CC	identified in human ECs. The human EC marker proteins and the		
CC	polynucleotide sequences encoding them are useful for detecting,		
CC	diagnosing or treating tumors as well as polycystic kidney disease,		
CC	diabetic retinopathy, rheumatoid arthritis, and psoriasis. They are also		
CC	useful for inhibiting neovascularization or tumour angiogenesis, for		
CC	inducing an immune response to tumour endothelial cells in a patient, or		
CC	for identifying candidate drugs for treating tumors. The present		
CC	sequence represents a human TEM or NEM gene of the invention		
XX			
SQ	Sequence 5540 BP; 1601 A; 1306 C; 1187 G; 1446 T; 0 U; 0 Other;		
Query Match	98.8%; Score 1090.8; DB 10; Length 5540;		
Best Local Similarity	99.8%; Pred. No. 0;		
Matches 1092; Conservative	0; Mismatches 2; Indels 0; Gaps 0		
QY	1 ATGGCCACGCGGAGCGAGAGAGCCCTCGGACATCGGCTTCAGTGGCTCTCTTTGGCCACT 60		
DB	144 ATGGCCACGCGGAGCGGAGAGAGCCCTCGGACATCGGCTTCAGTGGCTCTCTTTGGCCACT 203		
QY	61 CTGGTGGCTCATCTTGGCGCGGGAAGGGGGAGAGGAGGAGATGGGGGTCAGGCTGCTAC 120		
DB	204 CTGGTGGCTCATCTTGGCGCGGGAAGGGGGAGAGGAGGAGATGGGGGTCAGGCTGCTAC 263		

Qy	121	GGCGGATTTGACCTGTAACTTTCATTTTGGACAATACAGGAAGTGCTGCACACCTGGAT	180
Db	254	GGCGGATTTGACCTGTAACTTTCATTTTGGACAATACAGGAAGTGCTGCACACCTGGAT	323
Qy	181	GAATCTATTACTTTGTGGAAcAGTTGGCTCAcAAATTCATCAGCCCAcAGTTGGAATG	240
Db	324	GAATCTATTACTTTGTGGAAcAGTTGGCTCAcAAATTCATCAGCCCAcAGTTGGAATG	383
Qy	241	TCCTTTATTTGTTTTCTCCACCCGAGGAAcAAcCTTAATGAAcCTGCACAGAAcCAGAGA	300
Db	384	TCCTTTATTTGTTTTCTCCACCCGAGGAAcAAcCTTAATGAAcCTGCACAGAAcCAGAGA	443
Qy	301	CAAAATCCGTCAAAGGCTAGAAAGAAcCTCCAGAAAGTTCTGCAGAGAGAGAcCTTAACATG	360
Db	444	CAAAATCCGTCAAAGGCTAGAAAGAAcCTCCAGAAAGTTCTGCAGAGAGAGAcCTTAACATG	503
Qy	361	CATGAAAGATTTGAAAGGCGcAGTAGAGAcAGATTATATGAAACAGAcAAAGGATACAG	420
Db	504	CATGAAAGATTTGAAAGGCGcAGTAGAGAcAGATTATATGAAACAGAcAAAGGATACAG	563
Qy	421	ACAGCAACGTCATCCTATTGCTTTGACTGAATGAGAAcCTCCATGAAAGATCTCTTTTCTAT	480
Db	564	ACAGCAACGTCATCCTATTGCTTTGACTGAATGAGAAcCTCCATGAAAGATCTCTTTTCTAT	623
Qy	481	TCAGAGAGGAGGCTAATAGGCTGTGAGATCTTGGTGCAATTGTTACTGTGTGGTGTG	540
Db	624	TCAGAGAGGAGGCTAATAGGCTGTGAGATCTTGGTGCAATTGTTACTGTGTGGTGTG	683
Qy	541	AAAGATTTCAATGAGACAcAGCTGGCCCGGATTTGGGAcAGTAAGATCATGTGTTTCCC	600
Db	684	AAAGATTTCAATGAGACAcAGCTGGCCCGGATTTGGGAcAGTAAGATCATGTGTTTCCC	743
Qy	601	GTGAATGACGGCTTTCAGGCTCTGGAAGGCAcATCCACTGAATTTTGAAGAAcGCTGc	660
Db	744	GTGAATGACGGCTTTCAGGCTCTGGAAGGCAcATCCACTGAATTTTGAAGAAcGCTGc	803
Qy	661	ATCGAAATTTAGACAGCTGAAcCATACCAcCAATATGTGAGAGAGAGTCAATTTCAAGTTGC	720
Db	804	ATCGAAATTTAGACAGCTGAAcCATACCAcCAATATGTGAGAGAGAGTCAATTTCAAGTTGC	863
Qy	721	GTGAGAGAAAcCGGCTTCCGACATGCCGCAcGTGGAcAGGGTCTCTGCAGCTTCAAG	780
Db	864	GTGAGAGAAAcCGGCTTCCGACATGCCGCAcGTGGAcAGGGTCTCTGCAGCTTCAAG	923
Qy	781	ATCATATGACTGGGTGACACTGAATGAGAAGCCCTTTCTGTGGAAGACATTAATTACTG	840
Db	924	ATCATATGACTGGGTGACACTGAATGAGAAGCCCTTTCTGTGGAAGACATTAATTACTG	983
Qy	841	TGTCAGAGCCTATCTTAAAGAAAGTTGGACAGAAAGCTGCACTCAGCTCAGCATGAAc	900
Db	984	TGTCAGAGCCTATCTTAAAGAAAGTTGGACAGAAAGCTGCACTCAGCTCAGCATGAAc	1043
Qy	901	GATGGCCTCTCTTTTATCTCCAGTTCTGTcATCATCAcCAcCAcCAcCAcCTGTTCTGAcGT	960
Db	1044	GATGGCCTCTCTTTTATCTCCAGTTCTGTcATCATCAcCAcCAcCAcCAcCTGTTCTGAcGT	1103
Qy	961	TCcATcCTcGcGCATcGcCcCTGcTGAATcCTGTTCCGcTcCTTAAGCcCTGcGCTcCTcCTG	1020
Db	1104	TCcATcCTcGcGCATcGcCcCTGcTGAATcCTGTTCCGcTcCTTAAGCcCTGcGCTcCTcCTG	1163
Qy	1021	TGcGTTCTGcGcCcCTcCTGcTGAcCTGTGAATATCAAGAGAGTCCCTCAcCCcCTGcCGAG	1080
Db	1164	TGcGTTCTGcGcCcCTcCTGcTGAcCTGTGAATATCAAGAGAGTCCCTCAcCCcCTGcCGAG	1223
Qy	1081	GAGAGTGAAGGAAA 1094	
Db	1224	GAGAGTGAAGGAGA 1237	

00000

ADR48215
ID ADR48215 standard; cDNA; 5540 BP.
XX

AC	ADR48215;
XX	
DT	18-NOV-2004 (first entry)
XX	
DE	Human tumour endothelial marker 8 precursor encoding cDNA SEQ.3.
XX	
KM	pancreatic cancer-associated transcript; pancreatic cancer; human;
XX	cystostic; gene therapy; protein therapy;
KW	tumour endothelial marker 8 precursor; TEM8; gene; ss.
XX	
OS	Homo sapiens.
XX	
FT	Key Location/Qualifiers
FT	CDS 144..1838
FT	/tag= a
XX	/product= "tumour endothelial marker 8 precursor (TEM8) "
PN	WO2004074510-A1.
XX	
PD	02-SEP-2004.
XX	
PP	18-FEB-2004; 2004WO-AU000194.
XX	
PR	18-FEB-2003; 2003AU-00900747.
XX	
PA	(GARV-) GARVAN INST MEDICAL RES.
XX	
PI	Blankin A, Segara D, Henshall S, Sutherland R;
DR	WPI, 2004-635591/61.
XX	P-P8DB; ADR48216.
PT	Detecting pancreatic cancer-associated transcript in a biological sample,
XX	useful for diagnosing or treating the disease, comprises contacting the
PT	sample with a polynucleotide that selectively hybridises to a specific
XX	sequence.
PS	Claim 70; SEQ ID NO 3; 263bp; English.
XX	
CC	The present invention describes a method for detecting a pancreatic
CC	cancer-associated transcript in a biological sample. The method comprises
CC	contacting the biological sample with a polynucleotide that selectively
CC	hybridises to a sequence at least 80% identical to a sequence as shown in
CC	any one of Tables 3 to 25 in the specification or having the Genbank
CC	Accession Number AF279145. Also described: (1) diagnosing pancreatic
CC	cancer in a human or animal subject being tested, determining the
CC	likelihood that a subject having a pancreatic cancer will survive, or
CC	determining the suitability of a subject having a pancreatic cancer for
CC	surgical resection therapy; (2) detecting a pancreatic cancer-associated
CC	polypeptide in a biological sample; (3) determining the likelihood that a
CC	subject having a pancreatic cancer will survive; and (4) monitoring the
CC	efficacy of a therapeutic treatment of pancreatic cancer. A pancreatic
CC	cancer-associated transcript has cytostatic activity, and can be used in
CC	gene and protein therapy. A pancreatic cancer-associated transcript
CC	polynucleotide, a vector comprising the polynucleotide, an isolated
CC	polypeptide or an antibody that binds to the isolated polypeptide can be
CC	used for diagnosing or prognosing pancreatic cancer or for preparing a
CC	medicament for the treatment of pancreatic cancer. The prognostic or
CC	diagnostic methods are useful for the early detection of pancreatic
CC	cancer or its metastases, and for monitoring the progress of disease such
CC	as during remission or following surgery or chemotherapy. The present
CC	sequence encodes human tumour endothelial marker 8 precursor (TEM8),
CC	which is used in the exemplification of the present invention.
XX	
SQ	Sequence 5540 BP; 1601 A; 1306 C; 1187 G; 1446 T; 0 U; 0 Other;
	Query Match 98.8%; Score 1090.8; DB 13; Length 5540;
	Best Local Similarity 99.8%; Pred. No. 0;
	Matches 1092; Conservative 0; Mismatches 2; Indels 0; Gaps 0
GZ	1 ATGGCGACGCGGAGCGGAGCCTTGGGCATTCGACTTGCTCTTTGGCACT 60
DB	144 ATGGCGACGCGGAGCGGAGCCTTGGGCATTCGACTTGCTCTTTGGCACT 203

QY 61 CTGGTCTCATCTGCGCCGCGCAAGGGGAGCGCAGGGAGATGGGGGTCCAGCTGCTAC 120
 DB 204 CTGGTCTCATCTGCGCCGCGCAAGGGGAGCGCAGGGAGATGGGGGTCCAGCTGCTAC 263
 QY 121 GGGCGATTGACCTGATCTTCAATTTGGACAAATCAGAAAGTGTGCTGCACCTGGAAT 180
 DB 264 GGGCGATTGACCTGATCTTCAATTTGGACAAATCAGAAAGTGTGCTGCACCTGGAAT 323
 QY 181 GAATCTATTCTTTGTTGGAAACAGTGTGCTCAAAATTCATCAGCCCACTGTAGAAATG 240
 DB 324 GAATCTATTCTTTGTTGGAAACAGTGTGCTCAAAATTCATCAGCCCACTGTAGAAATG 383
 QY 241 TCTTTATTTGTTTCTCCACCCGAGAAACAATTAAGAACTGACAGAGACAGAGAA 300
 DB 384 TCTTTATTTGTTTCTCCACCCGAGAAACAATTAAGAACTGACAGAGACAGAGAA 443
 QY 301 CAAATCCGTCAAGGCTTGAAGAACTCCAGAAAGTTCTGCGAGAGAGACACTTACATG 360
 DB 444 CAAATCCGTCAAGGCTTGAAGAACTCCAGAAAGTTCTGCGAGAGAGACACTTACATG 503
 QY 361 CATAGAGATTGAAAGGCGCAGTGAAGATTTATTAAGAAACAGCAAGGGTTACAGG 420
 DB 504 CATAGAGATTGAAAGGCGCAGTGAAGATTTATTAAGAAACAGCAAGGGTTACAGG 563
 QY 421 ACAGCCAGGCTCATCATTTGCTTGTGACTGTAGAGAACTCATGAAGATCTTTTTCAT 480
 DB 564 ACAGCCAGGCTCATCATTTGCTTGTGACTGTAGAGAACTCATGAAGATCTTTTTCAT 623
 QY 481 TCAAGAGAGAGGCTTAATAGTCTCGAGATCTTGGTCAATTTGTTACTGTGTGTG 540
 DB 624 TCAAGAGAGAGGCTTAATAGTCTCGAGATCTTGGTCAATTTGTTACTGTGTGTG 683
 QY 541 AAAGATTCATGAGACACAGCTGGCCGGAATGCGGACAGTAAGATAGTGTTC 600
 DB 684 AAAGATTCATGAGACACAGCTGGCCGGAATGCGGACAGTAAGATAGTGTTC 743
 QY 601 GTGATGAGGCTTTCAGGCTCTGCAAGGATCATCTCACTCAATTTTGAAGAGTCTG 660
 DB 744 GTGATGAGGCTTTCAGGCTCTGCAAGGATCATCTCACTCAATTTTGAAGAGTCTG 803
 QY 661 ATCGAATTTCTAGCAGCTGAACCATCCACATATGTGAGAGAGTCAATTCAGGTTG 720
 DB 804 ATCGAATTTCTAGCAGCTGAACCATCCACATATGTGAGAGAGTCAATTCAGGTTG 863
 QY 721 GTGAGAGGAAAGGCTTCGACATGCGCGCAACGCGACAGGGTCTGCGAGCTTCAAG 780
 DB 864 GTGAGAGGAAAGGCTTCGACATGCGCGCAACGCGACAGGGTCTGCGAGCTTCAAG 923
 QY 781 ATCAATGACTCGTCACTCACTCAATGAGAAAGCCTTTCTGTGAAGACACTTATTTA 840
 DB 924 ATCAATGACTCGTCACTCACTCAATGAGAAAGCCTTTCTGTGAAGACACTTATTTA 983
 QY 841 TGTCCAGGCGCTATTTTAAAGAGTGTGCAATGAAGCTGCACTCAGGTACGATGAC 900
 DB 984 TGTCCAGGCGCTATTTTAAAGAGTGTGCAATGAAGCTGCACTCAGGTACGATGAC 1043
 QY 901 GATGGCCTCTTTTATCTCAGTCTGTCAATCATCACACACACACTGTTTGAAGGCT 960
 DB 1044 GATGGCCTCTTTTATCTCAGTCTGTCAATCATCACACACACACTGTTTGAAGGCT 1103
 QY 961 TCCATCTGTGCGCATGCGCCTGTGATCTGTCTCTGCTCAGCCCTGCTCTCTG 1020
 DB 1104 TCCATCTGTGCGCATGCGCCTGTGATCTGTCTCTGCTCAGCCCTGCTCTCTG 1163
 QY 1021 TGTGTTTGGCCCTCTGTGCTGCTGTATTCAGAGAGGTCCTCCACCCCTGCGAG 1080
 DB 1164 TGTGTTTGGCCCTCTGTGCTGCTGTATTCAGAGAGGTCCTCCACCCCTGCGAG 1223
 QY 1081 GAGAGTGAGAGAA 1094
 DB 1224 GAGAGTGAGAGAA 1237

RESULT 8
 AAD05303
 ID AAD05303 standard; cDNA; 2447 BP.
 XX
 AC AAD05303;
 XX
 DT 17-JUL-2001 (first entry)
 XX
 DE Human secreted protein-encoding gene 4 cDNA clone HMLFR02, SEQ ID NO:14.
 XX
 KW Human; secreted protein; proliferative disorder; cancer; tumour;
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
 KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
 KW cardiovascular disorder; angiogenic disorder; kidney disorder;
 KW gastrointestinal disorder; pregnancy-related disorder;
 KW endocrine disorder; infection; wound healing; vunerary; cell culture;
 KW chemotaxis; food additive; gene therapy; binding partner identification;
 KW chromosome 19; ss.
 KM
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 140..1351
 FT /*tag= a
 FT /product= "Human secreted protein"
 FT sig_peptide 140..220
 FT /*tag= b
 FT mat_peptide 221..1348
 FT /*tag= c
 FT /product= "Mature human secreted protein"
 FT
 XX
 PN W0200134626-A1.
 XX
 PD 17-MAY-2001.
 XX
 PE 01-NOV-2000; 2000MO-US030045.
 XX
 PF 05-NOV-1999; 99US-0163581P.
 PR 30-JUN-2000; 2000US-0215133P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX
 PI Ruben SM, Komatsoulis GA, Moore PA, Birse CE, Ni J;
 XX WPI; 2001-308778/32.
 DR P-PSDB; AAE01439.
 DR
 XX
 PT New nucleic acid molecules encoding 28 human secreted proteins for
 PT diagnosing, preventing, treating or ameliorating medical conditions and
 PT used as food additives or preservatives.
 PT
 PS Claim 1; Page 425-426; 562pp; English.
 XX
 XX AAD05300-AAD05379 represent cDNAs corresponding to 28 human secreted
 CC protein genes, and AAE01436-AAE01513 represent the proteins they encode.
 CC AAE01514-AAE01544 represent human secreted protein fragments or variants.
 CC The genes and their secreted proteins are useful for preventing, treating,
 CC or ameliorating medical conditions, e.g., by protein or gene therapy.
 CC Pathological conditions can be diagnosed by determining the amount of the
 CC new protein in a sample or by determining the presence of mutations in
 CC the new genes. Specific uses are described for each of the 28 genes,
 CC based on the tissues in which they are most highly expressed, and include
 CC developing products for the diagnosis or treatment of proliferative
 CC disorders, cancer, tumours, foetal and developmental abnormalities,
 CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
 CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,
 CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
 CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
 CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,

XX 24-JUN-2002; 2002US-00201292.
 PR 20-DEC-2001; 2001US-00038307.
 XX (ROTT/) ROTTMAN J B.
 PA (OKEE/) O'KEEFE T L.
 PA (OZKA/) OZKAYNAK E.
 PA (HEAL/) HEALEY J J.
 XX
 PI Rottman JB, O'Keefe TL, Ozkaynak E, Healey JJ;
 DR WPI; 2003-720708/68.
 DR P-PSDB; ADI00550.
 XX
 PT New TANGO 197 and/or TANGO 216 fusion polypeptides for preventing or
 PT ameliorating symptoms of anthrax comprises a von Willebrand factor A-like
 PT domain (vWF) amino acid sequence and an amino acid sequence heterologous
 PT to the vWF.
 XX
 PS Example; SEQ ID NO 17; 86bp; English.
 XX
 CC The invention relates to a novel fusion polypeptide comprising a von
 CC Willebrand factor A-like domain (vWF) amino acid sequence and an amino
 CC acid sequence heterologous to the vWF. The polypeptide of the invention
 CC demonstrates antibacterial activities whilst the composition and method
 CC may be useful in preventing or ameliorating the symptoms of cutaneous
 CC and/or inhalation anthrax. The current sequence is that of the human
 CC TANGO 197 Ig mutated fusion DNA of the invention.
 XX
 SQ Sequence 1674 BP; 433 A; 461 C; 424 G; 356 T; 0 U; 0 Other;
 Query Match 87.0%; Score 960.8; DB 10; Length 1674;
 Best Local Similarity 99.8%; Pred. No. 3.9e-284;
 Matches 962; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCACGGCGGAGAGAGCCCTCGCATCGGCTTCAGTGGCTCTCTTTGGCCACT 60
 DB 13 ATGGCCACGGCGGAGAGAGCCCTCGCATCGGCTTCAGTGGCTCTCTTTGGCCACT 72
 QY 61 CTGGTGTCTATCTGCGCGGCGGAGAGAGAGAGATGGGGTCCAGCTCTGCTAC 120
 DB 73 CTGGTGTCTATCTGCGCGGCGGAGAGAGAGAGATGGGGTCCAGCTCTGCTAC 132
 QY 121 GGGCGATTTGACCTGACTTCAATTTGGCAAAATCAGAGATGTCGACACACTGGAT 180
 DB 133 GGGCGATTTGACCTGACTTCAATTTGGCAAAATCAGAGATGTCGACACACTGGAT 192
 QY 181 GAAATCTATTACTTTGTGGAACAGTTGGCTCAAAATTCATCAGCCCAAGTTGAGATG 240
 DB 193 GAAATCTATTACTTTGTGGAACAGTTGGCTCAAAATTCATCAGCCCAAGTTGAGATG 252
 QY 241 TCCTTTATGTTTCTCCACCGAGAAACAATTGAAGTGAAGTGAAGAGAGAGAA 300
 DB 253 TCCTTTATGTTTCTCCACCGAGAAACAATTGAAGTGAAGTGAAGAGAGAGAA 312
 QY 301 CAAATCGGTCAAGGCTTGAAGAACTCCAGAAAGTTCGACGAGAGAGACATTTACATG 360
 DB 313 CAAATCGGTCAAGGCTTGAAGAACTCCAGAAAGTTCGACGAGAGAGACATTTACATG 372
 QY 361 CATGAAGATTTGAAGGGCCAGTGAAGAGATTTATATGAAAAAGAGAGAGATGAGG 420
 DB 373 CATGAAGATTTGAAGGGCCAGTGAAGAGATTTATATGAAAAAGAGAGAGATGAGG 432
 QY 421 ACAGCCAGGCTCATCTGCTTTGATGATGAGAGAACTCCATGAAGATCTTTTTCAT 480
 DB 433 ACAGCCAGGCTCATCTGCTTTGATGATGAGAGAACTCCATGAAGATCTTTTTCAT 492
 QY 481 TCAGAGAGGAGGCTTAATAGGTCGAGATCTTGGTGAATTTTACTGTTGGTGGTGG 540
 DB 493 TCAGAGAGGAGGCTTAATAGGTCGAGATCTTGGTGAATTTTACTGTTGGTGGTGG 552
 QY 541 AAGATTTCAATGAGACACAGCTGGCCGGAATTGCGGACAGTAAGATCATGTGTTCCC 600

DB 553 AAGATTTCAATGAGACACAGCTGCCCGGATTCGGAAGTAAGATCATGTGTTCCC 612
 QY 601 GTGAATGACGGCTTTCAGGCTCTGCAAGGATCATTCAGTCAATTTGAAGAAGCTTGC 660
 DB 613 GTGAATGACGGCTTTCAGGCTCTGCAAGGATCATTCAGTCAATTTGAAGAAGCTTGC 672
 QY 661 ATCGAAATTTAGAGAGTGAACCATCCACATATGTCAGAGAGATCATTTCAAGTTGTC 720
 DB 673 ATCGAAATTTAGAGAGTGAACCATCCACATATGTCAGAGAGATCATTTCAAGTTGTC 732
 QY 721 GTGAGAGAAAAGGCTTCCGACATGCGCGCAAGCTGACAGGGTCTCTGCAAGTTCAAG 780
 DB 733 GTGAGAGAAAAGGCTTCCGACATGCGCGCAAGCTGACAGGGTCTCTGCAAGTTCAAG 792
 QY 781 ATCAATAGCTGGGACACATCAATGAGAGAGCCCTTCTGAGAGACATTAATTACTG 840
 DB 793 ATCAATAGCTGGGACACATCAATGAGAGAGCCCTTCTGAGAGATTAATTACTG 852
 QY 841 TGTCACAGGCTTATCTTAAAGAAAGTTGGCATGAAAGCTGCACTCCAGTCCAGCATGAC 900
 DB 853 TGTCACAGGCTTATCTTAAAGAAAGTTGGCATGAAAGCTGCACTCCAGTCCAGCATGAC 912
 QY 901 GATGGCTCTCTTTTATCTCCAGTTGTCATCATCAACACACACTGTTCTGACGGT 960
 DB 913 GATGGCTCTCTTTTATCTCCAGTTGTCATCATCAACACACACTGTTCTGACGGT 972
 QY 961 TCCA 964
 DB 973 CCA 976

RESULT 11
 ADM64575
 ID ADM64575 standard; DNA, 1674 BP.
 XX
 AC ADM64575;
 XX
 DT 03-JUN-2004 (first entry)
 XX
 DB Mouse TANGO197-immunoglobulin (Ig) fusion protein DNA.
 XX
 XX antibacterial; gene therapy;
 KW von Willebrand factor A-like domain amino acid sequence;
 KW vWF amino acid sequence; anthrax; Bacillus anthracis; cutaneous anthrax;
 KW inhalation anthrax; mouse; TANGO197; immunoglobulin; Ig; fusion protein;
 KW gene; ds.
 XX
 OS Mus BD.
 OS Synthetic.
 OS
 PN US200134786-A1.
 XX
 PD 17-JUL-2003.
 XX
 PF 20-DEC-2001; 2001US-00038307.
 PR 20-DEC-2001; 2001US-00038307.
 XX
 PA (ROTT/) ROTTMAN J B.
 PA (OKEE/) O'KEEFE T L.
 PA (OZKA/) OZKAYNAK E.
 PA (HEAL/) HEALEY J J.
 XX
 PI Rottman JB, O'Keefe TL, Ozkaynak E, Healey JJ;
 DR WPI; 2003-829643/77.
 DR P-PSDB; ADM64576.
 XX
 PT New TANGO 197 and/or TANGO 216 fusion polypeptides for preventing or
 PT ameliorating symptoms of anthrax comprises a von Willebrand factor A-like
 PT domain (vWF) amino acid sequence and an amino acid sequence heterologous
 PT to the vWF.

XX Example; SEQ ID NO 17; 64pp; English.
PS The invention describes a fusion polypeptide comprising a von Willebrand
CC factor A-like domain (VWF) amino acid sequence and an amino acid sequence
CC heterologous to the VWF. Also described are: a method of preventing or
CC ameliorating a symptom of anthrax in a subject thought to be at risk for
CC exposure to or suspected of having been exposed to Bacillus anthracis;
CC and a pharmaceutical composition comprising the novel fusion polypeptide.
CC The composition and method are useful in preventing or ameliorating
CC symptoms of cutaneous and/or inhalation anthrax. This sequence encodes a
CC fusion protein comprising mouse TANCO197 and immunoglobulin (Ig) that can
CC be used to treat exposure to or prevent a symptom of anthrax.
XX
SQ Sequence 1674 BP; 433 A; 461 C; 424 G; 356 T; 0 U; 0 Other;

Query Match 87.0%; Score 960.8; DB 11; Length 1674;
Best Local Similarity 99.8%; Pred. No. 3.9e-284;
Matches 962; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCACGCGCGAGCGGAGAGCCCTCGGCATCGGCTTCCAGTGGCTCTCTTTGGCCACT 60
DB 13 ATGGCCACGCGCGAGCGGAGAGCCCTCGGCATCGGCTTCCAGTGGCTCTCTTTGGCCACT 72

QY 61 CTGGTCTCATCTGCGCCGCGCAAGGGGAGCGCAGGGAGATGGGGGTCCAGCTGCTAC 120
DB 73 CTGGTCTCATCTGCGCCGCGCAAGGGGAGCGCAGGGAGATGGGGGTCCAGCTGCTAC 132

QY 121 GGGGATTTGACCTGACTTCAATTTTGGCAAAATCAGGAAGTGGTGGCACTGCACTGGAAAT 180
DB 133 GGGGATTTGACCTGACTTCAATTTTGGCAAAATCAGGAAGTGGTGGCACTGCACTGGAAAT 192

QY 181 GAAATCTATTAATCTTTGTGGAACAGTTGGCTCAAAATTCATCAGCCCACTGTGAGATG 240
DB 193 GAAATCTATTAATCTTTGTGGAACAGTTGGCTCAAAATTCATCAGCCCACTGTGAGATG 252

QY 241 TCCCTTATGTTTCTTCCACCCGAGAGAAACCTTAATGAACCTGACAGAAAGAGAGAA 300
DB 253 TCCCTTATGTTTCTTCCACCCGAGAGAAACCTTAATGAACCTGACAGAAAGAGAGAA 312

QY 301 CAAATCGGTCAAGGCTGAGAGAACTCCAGAAAGTTCGCGAGGAGAGACACTTACATG 360
DB 313 CAAATCGGTCAAGGCTGAGAGAACTCCAGAAAGTTCGCGAGGAGAGACACTTACATG 372

QY 361 CATGAAGATTTGAAAGGCGCAGTGAGCAGATTTATATGAAAAAGAGAGGATACAGG 420
DB 373 CATGAAGATTTGAAAGGCGCAGTGAGCAGATTTATATGAAAAAGAGAGGATACAGG 432

QY 421 ACAGCCAGCGTCATCATTTGCTTTGATGATGAGAGAACTCCATGAAATCTCTTTTCTAT 480
DB 433 ACAGCCAGCGTCATCATTTGCTTTGATGATGAGAGAACTCCATGAAATCTCTTTTCTAT 492

QY 481 TCAGAGAGGAGGCTAATAGGCTCGAGATCTGGTGCAATTTGTTTACTGTTGGTGGTG 540
DB 493 TCAGAGAGGAGGCTAATAGGCTCGAGATCTGGTGCAATTTGTTTACTGTTGGTGGTG 552

QY 541 AAAAGATTTCAATGAGACACAGCTGGCCCGGATTTGCGGACAGTAAAGATCATGTGTTCC 600
DB 553 AAAAGATTTCAATGAGACACAGCTGGCCCGGATTTGCGGACAGTAAAGATCATGTGTTCC 612

QY 601 GTGAATGACGGCTTTCAGGCTCTGCAAGGCAATCATCCACTCAATTTTGAAGAAGTCCGCG 660
DB 613 GTGAATGACGGCTTTCAGGCTCTGCAAGGCAATCATCCACTCAATTTTGAAGAAGTCCGCG 672

QY 661 ATCGAAATTTCTAGAGCTGAACATCAATCATATGTGAGAGAGTCAATTTCAAGTTGTC 720
DB 673 ATCGAAATTTCTAGAGCTGAACATCAATCATATGTGAGAGAGTCAATTTCAAGTTGTC 732

QY 721 GTGAGAGAGAAACGGCTTCGACATGCGCGCAACGTGACAGGGTCTCTGACGCTTCAAG 780
DB 733 GTGAGAGAGAAACGGCTTCGACATGCGCGCAACGTGACAGGGTCTCTGACGCTTCAAG 792

QY 781 ATCAATGACTCGGTCACTCAATGAGAAAGCCCTTTTCTGTGAGAGACATTAATTTACTG 840

DB 793 ATCAATGACTCGGTCACTCAATGAGAAAGCCCTTTCTGTGAGAGATTAATTTACTG 852

QY 841 TGTCCAGCGGCTATCTTTAAAGAGTGGCATGAAAGCTGCACTCCAGGTCAAGATGAAAC 900
DB 853 TGTCCAGCGGCTATCTTTAAAGAGTGGCATGAAAGCTGCACTCCAGGTCAAGATGAAAC 912

QY 901 GATGGCTCTCTTTTATTTCTTCAGTCTGTGCATCATCAACCAACACTGTTTGACGGT 960
DB 913 GATGGCTCTCTTTTATTTCTTCAGTCTGTGCATCATCAACCAACACTGTTTGACGGT 972

QY 961 TCCA 964
DB 973 CCGA 976

RESULT 12
AD100545
ID AD100545 standard, DNA; 1650 BP.
XX
AC AD100545;
XX
DT 22-Apr-2004 (first entry)
XX
DE Human TANCO 197 Ig FcR mutated fusion DNA - plasmid pKTOK127.
XX
KW fusion; von Willebrand factor A-like domain; VWF; antibacterial;
XX cutaneous; inhalation anthrax; human; TANCO 197 Ig FcR fusion; mutant;
KW plasmid pKTOK127; ds; gene.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN US2003144193-A1.
XX
PD 31-Jul-2003.
XX
XX 24-Jul-2002; 2002US-00201292.
XX
PR 20-Dec-2001; 2001US-00038307.
XX
PA (ROTT/) ROTTMAN J B.
PA (OKEE/) O'KEEFE T L.
PA (OZKA/) OZKAYNAK E.
PA (HEAL/) HEALEY J J.
XX
PI Rottman JB, O'Keefe TL, Ozkaynak E, Healey JJ;
XX
DR WPI; 2003-720708/68.
DR P-PSDB; AD100546.
XX
PT New TANCO 197 and/or TANCO 216 fusion polypeptides for preventing or
PT ameliorating symptoms of anthrax comprises a von Willebrand factor A-like
PT domain (VWF) amino acid sequence and an amino acid sequence heterologous
PT to the VWF.
XX
XX
PS Example; SEQ ID NO 13; 86pp; English.
XX
XX The invention relates to a novel fusion polypeptide comprising a von
CC Willebrand factor A-like domain (VWF) amino acid sequence and an amino
CC acid sequence heterologous to the VWF. The polypeptide of the invention
CC demonstrates antibacterial activities whilst the composition and method
CC may be useful in preventing or ameliorating the symptoms of cutaneous
CC and/or inhalation anthrax. The current sequence is that of the human
CC TANCO 197 Ig FcR mutated fusion DNA of the invention.
XX
SQ Sequence 1650 BP; 427 A; 454 C; 419 G; 350 T; 0 U; 0 Other;

Query Match 86.3%; Score 952.4; DB 10; Length 1650;
Best Local Similarity 99.4%; Pred. No. 1.5e-281;
Matches 956; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 ATGGCCACGCGCGAGCGGAGAGCCCTCGGCATCGGCTTCCAGTGGCTCTCTTTGGCCACT 60

```

Db 1 ATGGCCAGGCGGAGGAGAGCCCTCGACATCGCTTCAGTGGCTCTCTTTGGCCACT 60
Qy 61 CTGGTGTCTCATCTGGCGCGGAGGAGGAGGAGGAGATGAGGAGTCCAGCTGTCTAC 120
Db 61 CTGGTGTCTCATCTGGCGCGGAGGAGGAGGAGGAGATGAGGAGTCCAGCTGTCTAC 120
Qy 121 GCGGAGATTGACCTGTACTTCAATTTTGGACAAATCAGGAGAGTGTCTGACCACTGGAAT 180
Db 121 GCGGAGATTGACCTGTACTTCAATTTTGGACAAATCAGGAGAGTGTCTGACCACTGGAAT 180
Qy 181 GAAATCTATTACTTTGTGGAAACGTTGGCTCACAATTCATCAGCCCAAGTTGAAATG 240
Db 181 GAAATCTATTACTTTGTGGAAACGTTGGCTCACAATTCATCAGCCCAAGTTGAAATG 240
Qy 241 TCCTTTATTTGTTTCTCACCCGAGAGAACACTTAATGAAACCTGACGAGAGAGAGAA 300
Db 241 TCCTTTATTTGTTTCTCACCCGAGAGAACACTTAATGAAACCTGACGAGAGAGAGAA 300
Qy 301 CAAATCCGCTCAAGGCTTGAAGAACTCCAGAAAGTTCTGCGAGAGAGAGACCTTACATG 360
Db 301 CAAATCCGCTCAAGGCTTGAAGAACTCCAGAAAGTTCTGCGAGAGAGAGACCTTACATG 360
Qy 361 CATGAAGATTGAAAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
Db 361 CATGAAGATTGAAAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
Qy 421 ACAGCAGCGTCATCATTTGCTTTGATGATGAGAGAACTCCAGAAAGTCTCTTTTTCAT 480
Db 421 ACAGCAGCGTCATCATTTGCTTTGATGATGAGAGAACTCCAGAAAGTCTCTTTTTCAT 480
Qy 481 TCAGAGAGGAGGAGGCTTAATAGGCTTCGAGATCTTGGTCAATTTGTTTACTGTGTGTG 540
Db 481 TCAGAGAGGAGGAGGCTTAATAGGCTTCGAGATCTTGGTCAATTTGTTTACTGTGTGTG 540
Qy 541 AAAAATTTCAATGAGACACAGCTGGCCGGAATGGGAGACAGTAAGATCATGTGTTTCCC 600
Db 541 AAAAATTTCAATGAGACACAGCTGGCCGGAATGGGAGACAGTAAGATCATGTGTTTCCC 600
Qy 601 GTGAATGACGCGCTTTCAGGCTCTGCAAGGAGCATCACTCAATTTTGAAGAAGCTCCGC 660
Db 601 GTGAATGACGCGCTTTCAGGCTCTGCAAGGAGCATCACTCAATTTTGAAGAAGCTCCGC 660
Qy 661 ATCGAAATTTCTAGAGAGTGAACCATCCACATATGTGAGAGAGATCATTTCAAGTGTG 720
Db 661 ATCGAAATTTCTAGAGAGTGAACCATCCACATATGTGAGAGAGATCATTTCAAGTGTG 720
Qy 721 GTGAGAGAGAAAGCGCTTCCGACATGCGCGCAACGTGACAGGAGTCTCTGACGCTTCAAG 780
Db 721 GTGAGAGAGAAAGCGCTTCCGACATGCGCGCAACGTGACAGGAGTCTCTGACGCTTCAAG 780
Qy 781 ATCAATGATCTCGGTCACATCAATGAGAAAGCCCTTTCTGTGAGAAACATTAATTAATG 840
Db 781 ATCAATGATCTCGGTCACATCAATGAGAAAGCCCTTTCTGTGAGAAACATTAATTAATG 840
Qy 841 TGTCACAGCGCTTATCTTAAGAAAGTGTGAGAAAGTGTGAGAAAGTGTGAGAAAGTGTG 900
Db 841 TGTCACAGCGCTTATCTTAAGAAAGTGTGAGAAAGTGTGAGAAAGTGTGAGAAAGTGTG 900
Qy 901 GATGGCTCTCTTTTATCTCAAGTTCTGTCAATCAACCAACCACTGTTTCCCAA 960
Db 901 GATGGCTCTCTTTTATCTCAAGTTCTGTCAATCAACCAACCACTGTTTCCCAA 960
Qy 961 TC 962
Db 961 TC 962

```

RESULT 13
 ID AD100547
 AC AD100547, standard; DNA, 1650 BP.

```

XX 22-Apr-2004 (first entry)
DT Human TANGO 197 Ig FcR WT fusion DNA - plasmid pUKTK129.
XX
DE fusion; von Willebrand factor A-like domain; vWF; antibacterial;
XX cutaneous; inhalation anthrax; human; TANGO 197 Ig FcR fusion; wild-type;
XX plasmid pUKTK129; ds; gene.
XX Homo sapiens.
XX US2003144193-A1.
XX
XX 31-JUL-2003.
XX
XX 24-JUL-2002; 2002US-00201292.
XX
XX 20-DEC-2001; 2001US-00038307.
XX
XX (ROTT/) ROTTMAN J B.
XX (OKEE/) O'KEEFE T L.
XX (OZKA/) OZKAYNAK E.
XX (HEAL/) HEALEY J J.
XX
XX Rottman JB, O'Keefe TL, Ozkaynak E, Healey JJ;
XX
XX WPI; 2003-720708/68.
XX P-PSDB; AD100548.
XX
XX New TANGO 197 and/or TANGO 216 fusion polypeptides for preventing or
XX ameliorating symptoms of anthrax comprises a von Willebrand factor A-like
XX domain (vWF) amino acid sequence and an amino acid sequence heterologous
XX to the vWF.
XX
XX Example; SEQ ID NO 15, 86pp; English.
XX
XX The invention relates to a novel fusion polypeptide comprising a von
XX Willebrand factor A-like domain (vWF) amino acid sequence and an amino
XX acid sequence heterologous to the vWF. The polypeptide of the invention
XX demonstrates antibacterial activities whilst the composition and method
XX may be useful in preventing or ameliorating the symptoms of cutaneous
XX and/or inhalation anthrax. The current sequence is that of the human
XX TANGO 197 Ig FcR wild-type fusion DNA of the invention.
XX
XX Sequence 1650 BP; 427 A; 453 C; 419 G; 351 T; 0 U; 0 Other;
XX
XX Query Match 86.3%; Score 952.4; DB 10; Length 1650;
XX Best Local Similarity 99.4%; Pred. No. 1.5e-281;
XX Matches 956; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
XX
Qy 1 ATGGCCAGGCGGAGGAGAGCCCTCGACATCGCTTCAGTGGCTCTCTTTGGCCACT 60
Db 1 ATGGCCAGGCGGAGGAGAGCCCTCGACATCGCTTCAGTGGCTCTCTTTGGCCACT 60
Qy 61 CTGGTGTCTCATCTGGCGCGGAGGAGGAGGAGGAGATGAGGAGTCCAGCTGTCTAC 120
Db 61 CTGGTGTCTCATCTGGCGCGGAGGAGGAGGAGGAGATGAGGAGTCCAGCTGTCTAC 120
Qy 121 GCGGAGATTGACCTGTACTTCAATTTTGGACAAATCAGGAGAGTGTCTGACCACTGGAAT 180
Db 121 GCGGAGATTGACCTGTACTTCAATTTTGGACAAATCAGGAGAGTGTCTGACCACTGGAAT 180
Qy 181 GAAATCTATTACTTTGTGGAAACGTTGGCTCACAATTCATCAGCCCAAGTTGAAATG 240
Db 181 GAAATCTATTACTTTGTGGAAACGTTGGCTCACAATTCATCAGCCCAAGTTGAAATG 240
Qy 241 TCCTTTATTTGTTTCTCACCCGAGAGAACCTTAATGAAACCTGACGAGAGAGAGAA 300
Db 241 TCCTTTATTTGTTTCTCACCCGAGAGAACCTTAATGAAACCTGACGAGAGAGAGAA 300
Qy 301 CAAATCCGCTCAAGGCTTGAAGAACTCCAGAAAGTTCTGCGAGAGAGAGACCTTACATG 360
Db 301 CAAATCCGCTCAAGGCTTGAAGAACTCCAGAAAGTTCTGCGAGAGAGAGACCTTACATG 360

```

QY 361 CATGAAGATTTTGAAGAGGCGAGTGAAGATTTATATGAAAACAGACAAAGGTACAGG 420
DB 361 CATGAAGATTTTGAAGAGGCGAGTGAAGATTTATATGAAAACAGACAAAGGTACAGG 420
QY 421 ACAGCCAGCGCTCATCTGCTTTGACTGATGAGAACTCCATGAAGATCTCTTTCTAT 480
DB 421 ACAGCCAGCGCTCATCTGCTTTGACTGATGAGAACTCCATGAAGATCTCTTTCTAT 480
QY 481 TCAGAGAGGAGGAGCTATAGGCTCGAGATCTGTATGATATGTTTCTGTGTGTG 540
DB 481 TCAGAGAGGAGGAGCTATAGGCTCGAGATCTGTATGATATGTTTCTGTGTGTG 540
QY 541 AAAGATTTCAATGAGACACAGCTGCGCGAGATTCGAGAGTGAAGATCTGTGTG 600
DB 541 AAAGATTTCAATGAGACACAGCTGCGCGAGATTCGAGAGTGAAGATCTGTGTG 600
QY 601 GTGAATGACGCGCTTTCAGGCTCTGCAAGGCTATCTCACTCAATTTTGAAGAAGTCTGC 660
DB 601 GTGAATGACGCGCTTTCAGGCTCTGCAAGGCTATCTCACTCAATTTTGAAGAAGTCTGC 660
QY 661 ATCGAAATTTAGAGAGCTGAACCATCCACCATATGTGAGAGAGTCAATTTCAAGTGTG 720
DB 661 ATCGAAATTTAGAGAGCTGAACCATCCACCATATGTGAGAGAGTCAATTTCAAGTGTG 720
QY 721 GTGAGAGGAAACGCGCTTCCGACATGCGCGCAACGTGACAGGCTCTGCAAGTGTG 780
DB 721 GTGAGAGGAAACGCGCTTCCGACATGCGCGCAACGTGACAGGCTCTGCAAGTGTG 780
QY 781 ATCAATGACTGCTGACACTCAATGAGAAAGCTTTTCTGTGAAAGACATTTACTG 840
DB 781 ATCAATGACTGCTGACACTCAATGAGAAAGCTTTTCTGTGAAAGACATTTACTG 840
QY 841 TGTCCAGGCGCTATCTTAAAGAAAGTGGCATGAAGCTGACCTCCAGTCAAGTGAAC 900
DB 841 TGTCCAGGCGCTATCTTAAAGAAAGTGGCATGAAGCTGACCTCCAGTCAAGTGAAC 900
QY 901 GATGCGCTCTTTTATCTCCAGTCTGTCTCATCATCAACACACACTGTTCTGACGGT 960
DB 901 GATGCGCTCTTTTATCTCCAGTCTGTCTCATCATCAACACACACTGTTCTCCAAA 960
QY 961 TC 962
DB 961 TC 962

RESULT 14
ADM64589
ID ADM64589 standard; DNA; 1650 BP.
XX
AC ADM64589;
XX
DT 03-JUN-2004 (first entry)
XX
DE TANGO197 extracellular domain-mutant IgG Fc fusion protein DNA #2.
XX
KW antibacterial; gene therapy;
XX von Willebrand factor A-like domain amino acid sequence;
KW vWF amino acid sequence; anthrax; Bacillus anthracis; cutaneous anthrax;
KW inhalation anthrax; human; TANGO197; Immunoglobulin G; IgG;
KW fragment of crystallisation; Fc; gene; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN US2003134786-A1.
XX
PD 17-JUL-2003.
XX
PF 20-DEC-2001; 2001US-00038307.
XX
PR 20-DEC-2001; 2001US-00038307.
XX

PA (ROT/) ROTTMAN J B.
PA (OKEE/) O'KEEFE T L.
PA (OZKA/) OZKAYNAK E.
PA (HEAL/) HEALEY J J.
PI Rottman JB, O'keefe TL, Ozkaynak E, Healey JJ;
XX
XX WPI; 2003-829643/77.
DR P-PSDB; ADM64590.
XX
XX New TANGO 197 and/or TANGO 216 fusion polypeptides for preventing or
PT ameliorating symptoms of anthrax comprises a von Willebrand factor A-like
PT domain (vWF) amino acid sequence and an amino acid sequence heterologous
PT to the vWF.
XX
XX Example; SEQ ID NO 13; 64bp; English.
XX
CC The invention describes a fusion polypeptide comprising a von Willebrand
CC factor A-like domain (vWF) amino acid sequence and an amino acid sequence
CC heterologous to the vWF. Also described are: a method of preventing or
CC ameliorating a symptom of anthrax in a subject thought to be at risk for
CC exposure to or suspected of having been exposed to Bacillus anthracis;
CC and a pharmaceutical composition comprising the novel fusion polypeptide.
CC The composition and method are useful in preventing or ameliorating
CC symptoms of cutaneous and/or inhalation anthrax. This sequence encodes a
CC fusion protein comprising human TANGO197 extracellular region, minus the
CC DG residues closest to the transmembrane region, and immunoglobulin G
CC (IgG) fragment of crystallisation (Fc) with mutations I235A and G237A.
XX
SQ Sequence 1650 BP; 427 A; 454 C; 419 G; 350 T; 0 U; 0 Other;
Query Match 86.3%; Score 952.4; DB 11; Length 1650;
Best Local Similarity 99.4%; Pred. No. 1.5e-281;
Matches 956; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 ATGGCCACGGCGGAGGAGAGGCGCTCGGATGGCTTCCAGTGGCTCTGTTGGCCACT 60
DB 1 ATGGCCACGGCGGAGGAGAGGCGCTCGGATGGCTTCCAGTGGCTCTGTTGGCCACT 60
QY 61 CTGGTGCTCATCTGCGCGCGGCAAGGGGAGCGAGGAGATGGGGTCCAGCTGTCTAC 120
DB 61 CTGGTGCTCATCTGCGCGCGGCAAGGGGAGCGAGGAGATGGGGTCCAGCTGTCTAC 120
QY 121 GCGGATTTGACCTGTACTTCTTGTGACAAATCAGAAAGTGTCTGCACTGTGAAT 180
DB 121 GCGGATTTGACCTGTACTTCTTGTGACAAATCAGAAAGTGTCTGCACTGTGAAT 180
QY 181 GAAATCTATTAATTGTGGAACAGTTGGCTCACAAAATTCATCGCCACAGTTGAGATG 240
DB 181 GAAATCTATTAATTGTGGAACAGTTGGCTCACAAAATTCATCGCCACAGTTGAGATG 240
QY 241 TCCCTTATGTTTCTCCACCGAGAAACACTTAATGAATGACAGAAACAGAGAA 300
DB 241 TCCCTTATGTTTCTCCACCGAGAAACACTTAATGAATGACAGAAACAGAGAA 300
QY 301 CAAATCCGTCAAGCGCTGAAGAACTTCAGAAAGTTTCCAGAGAGACACTTAATG 360
DB 301 CAAATCCGTCAAGCGCTGAAGAACTTCAGAAAGTTTCCAGAGAGACACTTAATG 360
QY 361 CAAATCCGTCAAGCGCTGAAGAACTTCAGAAAGTTTCCAGAGAGACACTTAATG 360
DB 361 CAAATCCGTCAAGCGCTGAAGAACTTCAGAAAGTTTCCAGAGAGACACTTAATG 360
QY 361 CATGAAGATTTTGAAGAGGCGAGTGAAGATTTATATGAAAACAGACAAAGGTACAGG 420
DB 361 CATGAAGATTTTGAAGAGGCGAGTGAAGATTTATATGAAAACAGACAAAGGTACAGG 420
QY 421 ACAGCCAGCGCTCATCTGCTTTGACTGATGAGAACTCCATGAAGATCTCTTTCTAT 480
DB 421 ACAGCCAGCGCTCATCTGCTTTGACTGATGAGAACTCCATGAAGATCTCTTTCTAT 480
QY 481 TCAGAGAGGAGGAGCTAATAGGCTCGAGATCTGTATGATATGTTTCTGTGTGTG 540
DB 481 TCAGAGAGGAGGAGCTAATAGGCTCGAGATCTGTATGATATGTTTCTGTGTGTG 540
QY 541 AAAGATTTCAATGAGACACAGCTGCGCGAGATTCGAGAGTGAAGATCTGTGTG 600
DB 541 AAAGATTTCAATGAGACACAGCTGCGCGAGATTCGAGAGTGAAGATCTGTGTG 600

Db 541 AAAAGTTTCATGACACAGCTGCGCCGGATTGCGGACAGTAAGATCATGTGTTCCC 600
Qy 601 GTGAATGACGGCTTTCAGGCTCTGCAAGCATCATCTCAATTTTGAAGAAGTCTGCG 660
Db 601 GTGAATGACGGCTTTCAGGCTCTGCAAGCATCATCTCAATTTTGAAGAAGTCTGCG 660
Qy 661 ATGGAATTCTAGACAGCTGAACCATCCACATATGTGACAGAGAGTCAATTTCAAGTTGTC 720
Db 661 ATGGAATTCTAGACAGCTGAACCATCCACATATGTGACAGAGAGTCAATTTCAAGTTGTC 720
Qy 721 GTGAGAGGAAACGGCTTCGACATGCGCGCAACGTCGACAGGTCCTGTCAGCTTCAG 780
Db 721 GTGAGAGGAAACGGCTTCGACATGCGCGCAACGTCGACAGGTCCTGTCAGCTTCAG 780
Qy 781 ATCAATGACTCGTCACACTCAATGAGAGGCTTTTCTGTGAAGACACTTATTTACTG 840
Db 781 ATCAATGACTCGTCACACTCAATGAGAGGCTTTTCTGTGAAGATCTTATTTACTG 840
Qy 841 TGTCCAGCGCTTATCTTAAAGAAAGTTGGCATGAAAGCTGCACTTCAGTCAGCATGAA 900
Db 841 TGTCCAGCGCTTATCTTAAAGAAAGTTGGCATGAAAGCTGCACTTCAGTCAGCATGAA 900
Qy 901 GATGCGCTCTTTTATCTCAGATTCTGTCTCATCATCACACACACTGTTCTGACGCT 960
Db 901 GATGCGCTCTTTTATCTCAGATTCTGTCTCATCATCACACACACTGTTCTGACGCT 960
Qy 961 TC 962
Db 961 TC 962

RESULT 15

ADM64591 ID ADM64591 standard; DNA; 1650 BP.

ADM64591;

03-JUN-2004 (first entry)

TANGO197 extracellular domain-wild type IgG Fc fusion protein DNA.

XX antibody; gene therapy;
XX von Willebrand factor A-like domain amino acid sequence;
XX vWF amino acid sequence; anthrax; Bacillus anthracis; cutaneous anthrax;
XX inhalation anthrax; human; TANGO197; immunoglobulin G; IgG;
XX fragment of crystallisation; Fc; gene; ds.

OS Homo sapiens.
OS Synthetic.

XX US2003134786-A1.

XX 17-JUL-2003.

XX 20-DEC-2001; 2001US-00038307.

XX 20-DEC-2001; 2001US-00038307.

XX (ROTT/) ROTTMAN J B.
XX (OKEE/) O'KEEFE T L.
XX (OZKA/) OZKAYNAK E.
XX (HEAL/) HEALEY J J.

XX Rottman JB, O'Keefe TL, Ozkaynak E, Healey JJ;

XX WPI; 2003-829643/77.

XX P-PSDB; ADM64592.

PT New TANGO 197 and/or TANGO 216 fusion polypeptides for preventing or
PT ameliorating symptoms of anthrax comprises a von Willebrand factor A-like
PT domain (vWF) amino acid sequence and an amino acid sequence heterologous
PT to the vWF.
XX

PS Example; SEQ ID NO 15; 64pp; English.

XX The invention describes a fusion polypeptide comprising a von Willebrand
CC factor A-like domain (vWF) amino acid sequence and an amino acid sequence
CC heterologous to the vWF. Also described are: a method of preventing or
CC ameliorating a symptom of anthrax in a subject thought to be at risk for
CC exposure to or suspected of having been exposed to Bacillus anthracis;
CC and a pharmaceutical composition comprising the novel fusion polypeptide.
CC The composition and method are useful in preventing or ameliorating
CC symptoms of cutaneous and/or inhalation anthrax. This sequence encodes a
CC fusion protein comprising human TANGO197 extracellular region, minus the
CC DG residues closest to the transmembrane region, and immunoglobulin G
CC (IgG) fragment of crystallisation (Fc).
XX

SQ Sequence 1650 BP; 427 A; 453 C; 419 G; 351 T; 0 U; 0 Other;

Query Match 86.3%; Score 952.4; DB 11; Length 1650;
Best Local Similarity 99.4%; Pred. No. 1.5e-281;
Matches 956; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 ATGSCACGCGGAGGAGAGCCCTGCGCATCGGCTTCCAGTGGCTCTTTGGCCACT 60
Db 1 ATGSCACGCGGAGGAGAGCCCTGCGCATCGGCTTCCAGTGGCTCTTTGGCCACT 60
Qy 61 CTGGTGTCTCATCTGCGCCGCGCAAGGCGGACGCGAGGAGATGCGGCTCTGCTTAC 120
Db 61 CTGGTGTCTCATCTGCGCCGCGCAAGGCGGACGCGAGGAGATGCGGCTCTGCTTAC 120
Qy 121 GCGGATTTGACCTGTACTTCTTATTTGGACAAATCAGAAAGTGTGTGACACCTGGAAT 180
Db 121 GCGGATTTGACCTGTACTTCTTATTTGGACAAATCAGAAAGTGTGTGACACCTGGAAT 180
Qy 181 GAAATCTTACTTGTGAGAACAGTTGGCTCACAAATTCATCAGCCACAGTTGAGATG 240
Db 181 GAAATCTTACTTGTGAGAACAGTTGGCTCACAAATTCATCAGCCACAGTTGAGATG 240
Qy 241 TCCCTTATTTGTTTCTCACCCGAGAACACTTAAATGAACTGACAGAAACAGAGAA 300
Db 241 TCCCTTATTTGTTTCTCACCCGAGAACACTTAAATGAACTGACAGAAACAGAGAA 300
Qy 301 CAAATCCGTCAAGGCTTGAAGAACTCCAGAAAGTTCTGCCAGAGAGACACTTAATG 360
Db 301 CAAATCCGTCAAGGCTTGAAGAACTCCAGAAAGTTCTGCCAGAGAGACACTTAATG 360
Qy 361 CATGAGATTTTGAAGAGGCGCAGTGAAGATTTATTTATGAAACAGACAGGTTACAG 420
Db 361 CATGAGATTTTGAAGAGGCGCAGTGAAGATTTATTTATGAAACAGACAGGTTACAG 420
Qy 421 ACAAGCAGCGTCATCTGCTTTGACTGATGAGAACTCCATGAAAGTCTTTTCTAT 480
Db 421 ACAAGCAGCGTCATCTGCTTTGACTGATGAGAACTCCATGAAAGTCTTTTCTAT 480
Qy 481 TCAGAGAGGAGGCTAATAGTCTCGAGATCTTGTCGAATTTGTTACTGTGTGTG 540
Db 481 TCAGAGAGGAGGCTAATAGTCTCGAGATCTTGTCGAATTTGTTACTGTGTGTG 540
Qy 541 AAAAGTTTCAATGAGACACAGCTGCGCGGATTTGCGGACGTAAAGATCATGTGTTCCC 600
Db 541 AAAAGTTTCAATGAGACACAGCTGCGCGGATTTGCGGACGTAAAGATCATGTGTTCCC 600
Qy 601 GTGAATGACGGCTTTCAGGCTCTGCAAGGATCATCCACTCAATTTTGAAGAAGTCTGCG 660
Db 601 GTGAATGACGGCTTTCAGGCTCTGCAAGGATCATCCACTCAATTTTGAAGAAGTCTGCG 660
Qy 661 ATGGAATTCTAGACAGCTGAACCATCCACATATGTGACAGAGAGTCAATTTCAAGTTGTC 720
Db 661 ATGGAATTCTAGACAGCTGAACCATCCACATATGTGACAGAGAGTCAATTTCAAGTTGTC 720
Qy 721 GTGAGAGGAAACGGCTTCGACATGCGCGCAACGTCGACAGGTCCTGTCAGCTTCAG 780
Db 721 GTGAGAGGAAACGGCTTCGACATGCGCGCAACGTCGACAGGTCCTGTCAGCTTCAG 780
Qy 781 ATCAATGACTCGTCACACTCAATGAGAGGCTTTTCTGTGAGAGACACTTATTTACTG 840

Db 781 ATCAATGACTCGGTCACTCACTCAATGAAAGCCCTTTCTGTGSAATTAATTACTGTG 840

QY 841 TGTCCAGCGCCCTATCTTTAAAGAAGTTGGCATGAAGCTGCACTCCAGGTCAGCATGAAC 900

Db 841 TGTCCAGCGCCCTATCTTTAAAGAAGTTGGCATGAAGCTGCACTCCAGGTCAGCATGAAC 900

QY 901 GATGGCCTCTCTTTTATCTCCAGTTCTGTTCATCATCACCACACACTGTTCTGACGGT 960

Db 901 GATGGCCTCTCTTTTATCTCCAGTTCTGTTCATCATCACCACACACTGTTCTCCCAA 960

QY 961 TC 962

Db 961 TC 962

Search completed: June 14, 2005, 14:06:24
Job time : 704 secs


```
QY 61 CTGGTGTCTATCTGCGCCGCGGGAAGGGAAGGAGAGATGGGGGTCCAGCCTCTAC 120
DB 369 CTGGGTCTATCTGCGCGGGAAGGGAAGGAGAGATGGGGGTCCAGCCTCTAC 428
QY 121 GCGCGATTTGACCTGTACTTCTATTTTGGCAATATCAGAAAGTGTCTGACCACTGGAAT 180
DB 429 GCGCGATTTGACCTGTACTTCTATTTTGGCAATATCAGAAAGTGTCTGACCACTGGAAT 488
QY 181 GAAATCTATTAATCTTTGTGGAAACAGTTGGCTCAAAATTCATCAGCCCAAGTTGAAATG 240
DB 489 GAAATCTATTAATCTTTGTGGAAACAGTTGGCTCAAAATTCATCAGCCCAAGTTGAAATG 548
QY 241 TCCCTTATTTGTTTCTCCACCGGAGGAACCTTAAATGAATGAATGAATGAATGAATG 300
DB 549 TCCCTTATTTGTTTCTCCACCGGAGGAACCTTAAATGAATGAATGAATGAATGAATG 608
QY 301 CAAATCCGTCAAGGCTTGAAGAACTCCAGAAAGTTCTGCGAGAGAGACACTTACATG 360
DB 609 CAAATCCGTCAAGGCTTGAAGAACTCCAGAAAGTTCTGCGAGAGAGACACTTACATG 668
QY 361 CATGAAGATTTGAAAAGGCGCAGTGAAGCATTTATTTATGAAAACAGCAAGGTTACAG 420
DB 669 CATGAAGATTTGAAAAGGCGCAGTGAAGCATTTATTTATGAAAACAGCAAGGTTACAG 728
QY 421 ACAGCCAGGTCATCATTTGCTTTGACTGATGAGAACTCCAGAAAGTCTCTTTTCTAT 480
DB 729 ACAGCCAGGTCATCATTTGCTTTGACTGATGAGAACTCCAGAAAGTCTCTTTTCTAT 788
QY 481 TCAGAGAGGAGGCTTAATAGTCTGAGATCTTGTGCAATGTTTACTGTGTGTGTG 540
DB 789 TCAGAGAGGAGGCTTAATAGTCTGAGATCTTGTGCAATGTTTACTGTGTGTGTGTG 848
QY 541 AAAGATTTCAATGAGACACAGCTGCGCGGATTTGGGACAGTAAGATCATGTGTTTCC 600
DB 849 AAAGATTTCAATGAGACACAGCTGCGCGGATTTGGGACAGTAAGATCATGTGTTTCC 908
QY 601 GTGAATGACGGCTTCAAGGCTCTGCAAGGATCATCACTCAATTTTGAAGAGCTCCG 660
DB 909 GTGAATGACGGCTTCAAGGCTCTGCAAGGATCATCACTCAATTTTGAAGAGCTCCG 968
QY 661 ATCGAATTTCTAGCAGCTGAACCATCAACATATGTGAGAGAGTCAATTTCAATGTC 720
DB 969 ATCGAATTTCTAGCAGCTGAACCATCAACATATGTGAGAGAGTCAATTTCAATGTC 1028
QY 721 GTGAGAGAAACGGCTTCCGACATGCCGCAACGTGACAGGGTCTCTGCAAGTTCAAG 780
DB 1029 GTGAGAGAAACGGCTTCCGACATGCCGCAACGTGACAGGGTCTCTGCAAGTTCAAG 1088
QY 781 ATCAATGACGCGGTCAACTCAATGAG 807
DB 1089 ATCAATGACGCGGTCAACTCAATGAG 1115

RESULT 2
US-09-799-451-250
; Sequence 250, Application US/09799451
; Patent No. 678369
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ma, Yungting
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wang, Dunhui
; APPLICANT: Yang, Yonghong
```

```
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Reena
; APPLICANT: Dimnac, Radoje T.
; TITLE OF INVENTION: No. 6783969e1 Nucleic Acids and
; FILE REFERENCE: 803
; CURRENT APPLICATION NUMBER: US/09/799,451
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: pc_fl_genes Version 2.0
; SEQ ID NO 250
; LENGTH: 3981
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (297)..(1118)
US-09-799-451-250

Query Match 46.8%; Score 516.8; DB 4; Length 3981;
Best Local Similarly 99.6%; Pred. No. 3.6e-159;
Matches 518; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 575 CGGACAGTAAGATCANGTGTTCCTCCGGAATGAACGGCTTCAAGCTCTGCAAGGCATCA 634
DB 1 CGGACAGTAAGATCANGTGTTCCTCCGGAATGAACGGCTTCAAGGCATCA 60
QY 635 TCCACTCAATTTTGAAGAGTCTGATTCGAATTTCTAGACGTGAACCATCACTAT 694
DB 61 TCCACTCAATTTTGAAGAGTCTGATTCGAATTTCTAGACGTGAACCATCACTAT 120
QY 695 GTGCAAGAGTCAATTTCAAGTGTCTGAGAGAAACGGCTTCCGACATCCGCAAG 754
DB 121 GTGCAAGAGTCAATTTCAAGTGTCTGAGAGAAACGGCTTCCGACATCCGCAAG 180
QY 755 TGGAACGGTCTCTGACGCTTCAAGATCAATGACTGCGTCACTCAATGAGAACCT 814
DB 181 TGGAACGGTCTCTGACGCTTCAAGATCAATGACTGCGTCACTCAATGAGAACCT 240
QY 815 TTTCTGTGAAGACATTAATTTACTGTGTCCAGGCGCTATCTTAAAGAAATGGCATGA 874
DB 241 TTTCTGTGAAGACATTAATTTACTGTGTCCAGGCGCTATCTTAAAGAAATGGCATGA 300
QY 875 AAGCTGACTCTCAGTGAAGCATGAACATGAGCTCTCTTTATCTCAATGTCATCA 934
DB 301 AAGCTGACTCTCAGTGAAGCATGAACATGAGCTCTCTTTATCTCAATGTCATCA 360
QY 935 TCACCAACACACATGTTCTGAGAGTTCATCTCGGACATGCGCTGTGATCTGTTCC 994
DB 361 TCACCAACACACATGTTCTGAGAGTTCATCTCGGACATGCGCTGTGATCTGTTCC 420
QY 995 TGCTCTAGCCCTGCTCTCTCTGTGTGTGAGCCCTCTGTGCACTGTGATTTATCA 1054
DB 421 TGCTCTAGCCCTGCTCTCTCTGTGTGTGAGCCCTCTGTGCACTGTGATTTATCA 480
QY 1055 AGGAGTTCCTCCACCCCTGCGAGAGAGTGAAGAAA 1094
DB 481 AGGAGTTCCTCCACCCCTGCGAGAGAGTGAAGAAA 520

RESULT 3
US-09-774-528-297
; Sequence 297, Application US/09774528
; Patent No. 6743619
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
```


Db 1011 CCATACGCTCAGAAATTAATTCGAGCTTGCTGAGGAGGAGCTTGGCAATYXX 1070
QY 749 GCAACGTGACAGGGTCTCTGACGCTTCAAGATCATGATGATGATGATGAGA 808
Db 1071 YY 1130
QY 809 AGCCCTTTCTGTGAGAACACTTATTACTGTGTCCAGGCGCTATCTTAAAGAACTTG 868
Db 1131 YY 1190
QY 869 GCATGAAGCTGACCTCAGGTACGATGAACATGCGCTCTTTTATCTCCAGTTCTG 928
Db 1191 YY 1250
QY 929 TCATCATCACACACACACACTGTCTGACGCTTCCATCTGCGCATGCGCTGCTGATCC 988
Db 1251 YY 1310
QY 989 TGTTCCTGCTCTGACGCTGCTCTCTCTGTGTGTCTGCGCCCTCTGCTGACGTGTA 1048
Db 1311 YY 1370
QY 1049 TTATCAAGAGGTCCTCCACCCCTGCTCC 1077
Db 1371 YYYYYYYYYYYYYYYYYYYYYYYYYYYY 1399

RESULT 5
US-09-621-976-8976/c
Sequence 8976, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J. B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J. Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621.976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 8976
LENGTH: 399
TYPE: DNA
ORGANISM: Homo sapiens
US-09-621-976-8976

Query Match 3.8%, Score 42.4, DB 4, Length 399;
Best Local Similarity 8.9%, Pred. No. 0.0031;
Matches 25, Conservative 142, Mismatches 113, Indels 0, Gaps 0,
QY 6 CACGCGAGAGGAGAGCCCTCGGATCGCTTCCAGTGTCTCTTTGGCACTGTGT 65
Db 284 SRGSGTRMGRYRMMGSGKSGWGGSYRMAGYRSMRSMYAMMKKMTCKKSSW 225
QY 66 GCTCATCTGCGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 125
Db 224 GSHSTGYAMMYKKSCTSRKMYKRRKRKRKCTSTKRTGRTGRTGRTGRTGRTG 165
QY 126 ATTTGACCTGTACTTCTTTTGCACAAATGAGAGGTGCTGTGACACATGGAAT 165
Db 164 RMTWRMYYKYSKWSKKTWRMTAYWTRMTKTRMTCTWCTWCTWCTWCTWCTW 105
QY 186 CTATATTCTTGAGAACAGTTGCTCAAAATTCATCAGCCACAGTTGAGAAATG 245
Db 104 RYRMYVYARAKRSKCTSTCTCYKMYAKKCKMSWMSMMKMSMMKMYKMY 45
QY 246 TATTTCTTCCACCGGAGAACACCTTAATGAACTG 285
Db 44 YMMKMSHMTYMSMSCYARKCMRTYATYTWTCMTWGR 5

RESULT 6

US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: VA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-F18
US-08-232-463-14

Query Match 3.8%, Score 41.4, DB 1, Length 7218;
Best Local Similarity 6.8%, Pred. No. 0.055;
Matches 27, Conservative 198, Mismatches 174, Indels 0, Gaps 0,
QY 263 GAGAGAACCTTATGAATGACAGAGACAGAGAACAAATCCGTCAAGCCTGAAG 322
Db 1407 RRR 1348
QY 323 AACTCGAAGAGTTCTGCGAGAGAGACACTTACATGATGAAGATTGAAGGCCA 382
Db 1347 RRR 1288
QY 383 GTGACAGATTATTTATGAAAACAGACAGAGGATGACAGACAGCAGCTATGCTT 442
Db 1287 RRR 1228
QY 443 TGACATGAGAGAACCTCAGAGATCTCTTTTCTATTCAGAGAGAGGCTAATAGT 502
Db 1227 RRR 1168
QY 503 CTCGACATTTGGTGCAATGTTTACTGTGTGTGTGAAGATTTCATGAGACAGC 562
Db 1167 RRR 1108
QY 563 TGCCCGGATGCGGACAGTAAAGATCATGTGTTCCGTAATGACGCTTCAGGCTC 622


```

? LENGTH: 1141
? TYPE: DNA
? ORGANISM: Artificial sequence
? FEATURE:
? NAME/KEY: Promoter
? LOCATION: (1)..(1141)
? OTHER INFORMATION: consensus sequence of A.c., L.a., and B.n. PAB1 promoters
?S-09-806-708B-22

```

Query Match	3.4%	Score 37.6;	DB 4;	Length 1141;
Best Local Similarity	10.9%;	Pred. No. 0.26;		
Matches	89;	Conservative 281;	Mismatches 443;	Indels 6;
			Gaps 1;	

QY	286	ACAGAGAAGACAGAAAGACCGGTCAAGGCGTGAAGAAAGCTCAAGAAAGTTGCGCAGAA	345
Db	987	AAATKARAGAGNNMRTYGAAAGKMGCMMAAATMBGWAADTAGKMCNNNNNNNTTIDVRMA	928
QY	346	GGAGACACTTACATGACATGAAGATTGAAAAGGCGCAGTAGACAGATTATATGAAGAAC	405
Db	927	MKAKNNNNNNNYWTFACYNRAATTNNKCAVTHMMKMTGHAHSKRTTRHHTTCRTTKYNNNNN	868
QY	406	AGACAAGGATACAGACACAGCCAGCTCATCTGTTGATCTGATGAGGAAGCTCCATGAA	465
Db	867	NARTVYVYHHAARRMMNMAWMTRTNNNNNNNNNACRNTRTWABWKHSWNNNNNNNNN	808
QY	466	GATCTCTTTTCTATTTCAGAGGAGGAGCTAATAGCTGCGAGACTTGCGTCAATGTT	525
Db	807	NNNNTWCHYTTANMBBCYRANNNNNAAAMARICNNYMHAAVTTHTDWCYKTMNTTWYMD	748
QY	526	TACTGTGTGCTGTGAAGAAATTCATGAGACACAGCTGGCCCGGATTTGCGGACAGTAAG	585
Db	747	MATTMBTTTTNMTTSTJMTNNNNNNMMACTNNNNNNMKAYAAHATNMGCGWNNNTDAR	688
QY	586	GATCATGTGTTCCCGTGATAGACGGCTTTGAGGCTGTGCAAGCATCATCCACTCAAT	645
Db	687	RTNNTTWRRRRMNTTKTRWISTTRHHHTGATNNNNNNNNNNNNNNNSCTCTRRMTM	628
QY	646	TTGAAGAAGTCCGTGATGAAATTTTACAGAGCTGAACCATCCACATATGTGACAGAGAG	705
Db	627	RWTMKGDMTVRKYKMKMDTTCTYDVWADSVVMVYVAYMMNRGRDVTYTRNNTYCKSVASH	568
QY	706	TCAATTTCAAGTGTGCTGAGAGAAAGCGCTTCGACATGCCCGCAAGTGAACAGGCTC	765
Db	567	YMYNSNNAMWRRRSARMSWMAWRTTNNNNMSGBVBRMAGTWMBRRNNNNNTDTRY	508
QY	766	CTGCGAGCTTCAAGATCAATGATGCTCGGTCACATCAATGAGAAAGCCCTTTCTGTGAA	825
Db	507	WWWKRWABITTVYDMSCKAKSMRGNNRRAMKMMWAAANDGANDHWTYMGNNNTMMNR	448
QY	826	GACACTTATTTACTGTGTCCAGCGGCTATCTTTAAAGAAAGTTGAGCAAGACCTGACCTC	885
Db	447	RAMKMMNNAMCBRAVCNNNNNNBAQVMHKKH-----MMRWTKYMKWKAQNNNNBKANY	394
QY	886	CAGGTGACGATGAAGACGATGGCGCTCTCTTTATCTCCAGTTCTGTATCATATCACACACAA	945
Db	393	MRVAMMWSRDTJNTDMMMWMTSDWEMHWYVDVYTWBRAMNNNNNNNNNNMBCKTSSWMMWD	334
QY	946	CACGTGTTGAAGGTTTCATCTCGGCACATCGCGCTGTGATCTGTGCTCGTCTCTAGCC	1005
Db	333	HMNTHCTYGNNTWSGAYEYMAAASMMAAAGASNBVTYNNCKRMTYMGKTYNNNNNNKANY	274
QY	1006	CTGGCTCTCTGTGTTGTTGCGGCCCTCTGTGCTGCACTGTGATTATCAAGAGGCTCCT	1065
Db	273	RTKTVAMCNRYYDYDTAVMTBKRNKYCYAVBYMYBMVYMGKHMBMWRABHRSNNMMVY	214
QY	1066	CCACCCCTCGCGGAGAGAGTGAAGAAATAAATAATAAAA	1104
Db	213	KCRNKYVWSMHYHARBYBKAAAVGCNNNNWCKDRAHHH	175

RESULT 10
US-08-728-323A-1/c
; Sequence 1, Application US/08728323A

Patent No. 5948676
GENERAL INFORMATION:
APPLICANT: Chang, Yuan
APPLICANT: Bohenzky, Roy A.
APPLICANT: Russo, James J.
APPLICANT: Edelman, Isidore S.
APPLICANT: Moore, Patrick S.
TITLE OF INVENTION: Immediate Early Protein From Kaposi's
TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA
ENCODING: Encoding Same And Uses Thereof
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Releasee #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/728,323A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/52268/JFW/MSC/SKS
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-378-0400
TELEFAX: 212-391-0525
INFORMATION FOR SEQ. ID NO.: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3489 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3489
US-08-728-323A-1

Query Match	3.4%	Score 37.4;	DB 2;	length 3489;
Best Local Similarity	50.3%;	Pred. No. 0.67;		
Matches	92;	Conservative	0;	Mismatches 91;
			Indels	0;
			Gaps	0;

Oy	883	CTCCAGGCTGACAGTAAAGCATGGCCCTCTCTTTATATCTCAAGTTCTGTCATATCAACACC	942
Db	1149	CTCCTCGTCATCCCTCCCTGTCATCTCTCTGTAATCTCTCCGTCATCTCTCTGTAATC	1096
Oy	943	ACACACGTGTTCTGACGGTTCATCTGAGCATGCGCCCTGTCGATCTGTTCTCTGCTCTTA	1002
Db	1089	CTCTCTGTCATACCTCTCTGTCATCTCTCTGTCATACCTCTCTGTCATCTCTCTGTCATC	1030
Oy	1003	GGCCCTGGCTCTCTCTGTGAGTTCTGGCCCTCTGCTGACACTGATTAATCAAGAGTTC	1062
Db	1029	CTCTCTGCTCCCTCATATCTGTCTCTGCTCTCTCTCTCATATCACTTAATGTATTGTCTATC	970
Oy	1063	CCT	1065
Db	969	CTT	967

RESULT 11
US-09-298-568-1/c
; Sequence 1, Application US/09298568
; Patent No. 6322792
; GENERAL INFORMATION:
; APPLICANT: Kieff, Elliott D.

APPLICANT: Ballester, Mary E.
APPLICANT: Kaye, Kenneth M.
TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO
FILE OF INVENTION: VIRUS DNA TO MEDIATE EFFICIENT EPISOME PERSISTENCE
FILE REFERENCE: 16412-10001R
CURRENT APPLICATION NUMBER: US/09/298,568
EARLIER APPLICATION NUMBER: US/09-04-21
EARLIER FILING DATE: 1998-11-19
NUMBER OF SEQ ID NOS: 3
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 1
LENGTH: 3489
TYPE: DNA
ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-298-568-1

Query Match 3.4%; Score 37.4; DB 3; Length 3489;
Best Local Similarity 50.3%; Pred. No. 0.67;
Matches 92; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

883 CTCGAGTCAGACATGAGGCTCTTTATCTCCAGTTCGTATCATACACACC 942
1149 CTCCTCGATCCTCTCTGTATCTCTCTGATCTCTCTGATCTCTCTGATC 1090
943 ACACACTGTTCTGACGGTTCATCTGACCATCGCCCTGATCTGTTCTGCTCTA 1002
1089 CTCCTCGATCCTCTCTGTATCTCTCTGATCTCTCTGATCTCTCTGATC 1030
1003 GCCCTGCTCTCTCTGTGTTCTGAGCCCTCTGCTGATCTGATTTAAGAGGTC 1062
1029 CTCCTGCTCTCTCTGTATCTCTCTGCTCTCTGATCTCTCTGATCTCTGATC 970

Qy 1063 CCT 1065
Db 969 CTT 967

RESULT 12
US-09-410-399-1/c
Sequence 1, Application US/09410399
Patent No. 6482587
GENERAL INFORMATION:
APPLICANT: Robertson, Erle S.
APPLICANT: Collier, Murray A.
TITLE OF INVENTION: Methods to Inhibit or Enhance the Binding of Viral DNA
FILE REFERENCE: US-03778
CURRENT APPLICATION NUMBER: US/09/410,399
CURRENT FILING DATE: 1999-10-01
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 1
LENGTH: 3489
TYPE: DNA
ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-410-399-1

Query Match 3.4%; Score 37.4; DB 4; Length 3489;
Best Local Similarity 50.3%; Pred. No. 0.67;
Matches 92; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

883 CTCGAGTCAGACATGAGGCTCTTTATCTCCAGTTCGTATCATACACACC 942
1149 CTCCTCGATCCTCTCTGTATCTCTCTGATCTCTCTGATCTCTCTGATC 1090
943 ACACACTGTTCTGACGGTTCATCTGACCATCGCCCTGATCTGTTCTGCTCTA 1002
1089 CTCCTCGATCCTCTCTGTATCTCTCTGATCTCTCTGATCTCTCTGATC 1030
1003 GCCCTGCTCTCTCTGTGTTCTGAGCCCTCTGCTGATCTGATTTAAGAGGTC 1062
1029 CTCCTGCTCTCTCTGTATCTCTCTGCTCTCTGATCTCTCTGATCTCTGATC 970

Qy 1063 CCT 1065
Db 969 CTT 967

RESULT 13
US-09-894-273-1/c
Sequence 1, Application US/09894273
Patent No. 6756203
GENERAL INFORMATION:
APPLICANT: Kieff, Elliott D.
APPLICANT: Ballester, Mary E.
TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO
FILE OF INVENTION: VIRUS DNA TO MEDIATE EFFICIENT EPISOME PERSISTENCE
FILE REFERENCE: 16412-10001R
CURRENT APPLICATION NUMBER: US/09/894,273
CURRENT FILING DATE: 2001-06-28
PRIOR APPLICATION NUMBER: US 60/109,422
PRIOR FILING DATE: 1998-11-19
NUMBER OF SEQ ID NOS: 3
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 1
LENGTH: 3489
TYPE: DNA
ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-894-273-1

Query Match 3.4%; Score 37.4; DB 4; Length 3489;
Best Local Similarity 50.3%; Pred. No. 0.67;
Matches 92; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

883 CTCGAGTCAGACATGAGGCTCTTTATCTCCAGTTCGTATCATACACACC 942
1149 CTCCTCGATCCTCTCTGTATCTCTCTGATCTCTCTGATCTCTCTGATC 1090
943 ACACACTGTTCTGACGGTTCATCTGACCATCGCCCTGATCTGTTCTGCTCTA 1002
1089 CTCCTCGATCCTCTCTGTATCTCTCTGATCTCTCTGATCTCTCTGATC 1030
1003 GCCCTGCTCTCTCTGTGTTCTGAGCCCTCTGCTGATCTGATTTAAGAGGTC 1062
1029 CTCCTGCTCTCTCTGTATCTCTCTGCTCTCTGATCTCTCTGATCTCTGATC 970

Qy 1063 CCT 1065
Db 969 CTT 967

RESULT 14
US-08-770-379-20
Sequence 20, Application US/08770379
Patent No. 5849564
GENERAL INFORMATION:
APPLICANT: Chang, Yuan
APPLICANT: Bohenzky, Roy A.
APPLICANT: Ruoso, James J.
APPLICANT: Edelman, Isidore S.
TITLE OF INVENTION: POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2005, 15:32:23 ; Search time 783 Seconds
(without alignments)
8740.438 Million cell updates/sec

Title: US-09-970-076-1_COPY_104_1207
Perfect score: 1104
Sequence: 1 atggccacggcgagcgag.....gtcgaagaaataaataaaa 1104

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 6046767 seqs, 3099530249 residues

Total number of hits satisfying chosen parameters: 12093534

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
- 20: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
- 21: /cgn2_6/ptodata/2/pubpna/US10J_PUBCOMB.seq:*
- 22: /cgn2_6/ptodata/2/pubpna/US10K_PUBCOMB.seq:*
- 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
- 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1104	100.0	1454	17	US-10-133-937-58 Sequence 58, Appl
2	1104	100.0	1454	17	US-10-159-563-58 Sequence 58, Appl
3	1090.8	98.8	5540	10	US-09-918-715-176 Sequence 176, Appl
4	1090.8	98.8	5540	10	US-09-918-715-231 Sequence 231, Appl
5	1090.8	98.8	5540	15	US-10-301-822-198 Sequence 198, Appl
6	1090.8	98.8	5540	20	US-10-474-794-176 Sequence 176, Appl
7	1090.8	98.8	5540	20	US-10-474-794-231 Sequence 231, Appl

	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45		
	960.8	960.8	960.8	960.8	960.8	960.8	960.8	960.8	960.8	960.8	960.8	960.8	960.8	960.8	960.8	960.8	960.8	960.8	960.8	960.8	960.8	960.8	960.8	960.8	960.8	960.8	960.8	960.8	960.8	960.8	960.8	960.8	960.8	960.8	960.8	960.8	960.8	960.8	960.8	960.8
	87.0	87.0	87.0	87.0	87.0	87.0	87.0	87.0	87.0	87.0	87.0	87.0	87.0	87.0	87.0	87.0	87.0	87.0	87.0	87.0	87.0	87.0	87.0	87.0	87.0	87.0	87.0	87.0	87.0	87.0	87.0	87.0	87.0	87.0	87.0	87.0	87.0	87.0	87.0	87.0
	1674	1674	1674	1674	1674	1674	1674	1674	1674	1674	1674	1674	1674	1674	1674	1674	1674	1674	1674	1674	1674	1674	1674	1674	1674	1674	1674	1674	1674	1674	1674	1674	1674	1674	1674	1674	1674	1674	1674	1674
	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	US-10-038-307-17	
	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	Sequence 17, Appl	

ALIGNMENTS

RESULT 1
US-10-133-937-58
Sequence 58, Application US/10133937
Publication NO. US20030207278A1
GENERAL INFORMATION:
APPLICANT: Khan, Javed
APPLICANT: Ringner, Markus
APPLICANT: Peterson, Carsten
APPLICANT: Melitzer, Paul
TITLE OF INVENTION: METHODS FOR ANALYZING HIGH DIMENSIONAL DATA FOR CLASSIFYING, DIAGNOSING, PROGNOSTICATING, AND/OR PREDICTING DISEASES AND
FILE REFERENCE: 11613.56US01
CURRENT APPLICATION NUMBER: US/10/133,937
NUMBER OF SEQ ID NOS: 99
SOFTWARE: PatentIn version 3.1
SEQ ID NO 58
LENGTH: 1454
TYPE: DNA
ORGANISM: Homo sapiens
US-10-133-937-58

Query Match 100.0%; Score 1104; DB 17; Length 1454;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 144 ATGGCCACGGCGAGCGAGAGCCCTTCGCGATCGCTTCAGTGGCTCTTTGGCCACT 60
1 ATGGCCACGGCGAGCGAGAGCCCTTCGCGATCGCTTCAGTGGCTCTTTGGCCACT 60
144 ATGGCCACGGCGAGCGAGAGCCCTTCGCGATCGCTTCAGTGGCTCTTTGGCCACT 203

QY 661 ATCGAAATTTAGAGCTGAAACCATTCACATATGTGAGAGAGTCAATTTCAAGTTGTC 720
 DB 804 ATCGAAATTTAGAGCTGAAACCATTCACATATGTGAGAGAGTCAATTTCAAGTTGTC 863
 QY 721 GTGAGAGAAACGGCTTCCGACATGCGGCAAGGTGACAGGGCTCTGCAAGTTCAAG 780
 DB 864 GTGAGAGAAACGGCTTCCGACATGCGGCAAGGTGACAGGGCTCTGCAAGTTCAAG 923
 QY 781 ATCAATGACTCGGTCACTCAATGAGAAAGCCCTTTCTGTGAGAACACTTATTTACTG 840
 DB 924 ATCAATGACTCGGTCACTCAATGAGAAAGCCCTTTCTGTGAGAACACTTATTTACTG 983
 QY 841 TGTCCAGCGCTTATCTTAAAGAAAGTTGGCATGAAAGTTCGACACTCCAGTTCAGATGAC 900
 DB 984 TGTCCAGCGCTTATCTTAAAGAAAGTTGGCATGAAAGTTCGACACTCCAGTTCAGATGAC 1043
 QY 901 GATGGCCCTCTTTATCTCAAGTTCTGATATCAATCAACACACACTGTTCTGAAGGT 960
 DB 1044 GATGGCCCTCTTTATCTCAAGTTCTGATATCAATCAACACACACTGTTCTGAAGGT 1103
 QY 961 TCCATCTGCGCAATCGCCCTGCTGATCTGTTCTGCTGAGCCCTGCTCTCTG 1020
 DB 1104 TCCATCTGCGCAATCGCCCTGCTGATCTGTTCTGCTGAGCCCTGCTCTCTG 1163
 QY 1021 TGGTTCTGCGCCCTCTGCTGCACTGTGATTATCAAGAGGTCCCTCCACCCCTGCGAG 1080
 DB 1164 TGGTTCTGCGCCCTCTGCTGCACTGTGATTATCAAGAGGTCCCTCCACCCCTGCGAG 1223
 QY 1081 GAGAGTGAGGAAATTAATTAATAA 1104
 DB 1224 GAGAGTGAGGAAATTAATTAATAA 1247

RESULT 3

US-09-918-715-176
 ; Sequence 176, Application US/09918715
 ; Publication No. US20030017157A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Brad St. Croix
 ; APPLICANT: Bert Vogelstein
 ; TITLE OF INVENTION: ENDOTHELIAL CBLL EXPRESSION PATTERNS
 ; FILE REFERENCE: 1107.00134
 ; CURRENT APPLICATION NUMBER: US/09/918, 715
 ; PRIOR APPLICATION NUMBER: 60/222,599
 ; PRIOR FILING DATE: 2000-08-02
 ; PRIOR APPLICATION NUMBER: 60/224,360
 ; PRIOR FILING DATE: 2000-08-11
 ; PRIOR APPLICATION NUMBER: 60/282,850
 ; PRIOR FILING DATE: 2000-04-11
 ; NUMBER OF SEQ ID NOS: 358
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 176
 ; LENGTH: 5540
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-918-715-176

Query Match 98.8%; Score 1090.8; DB 10; Length 5540;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1092; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCACGGCGGAGGAGGAGAGCCCTCGGATCGGCTTCAAGTGGCTCTTTGGCACT 60
 DB 144 ATGGCCACGGCGGAGGAGGAGAGCCCTCGGATCGGCTTCAAGTGGCTCTTTGGCACT 203
 QY 61 CTGGTGTCTATCTGCGCGGCAAGGGGAGCGAGGAGAGTGGGGTCCAGCTGCTAC 120
 DB 204 CTGGTGTCTATCTGCGCGGCAAGGGGAGCGAGGAGAGTGGGGTCCAGCTGCTAC 263
 QY 121 GGGGATTTGACTGTACTTCAATTTTGGACAAATCAGGAAGTGTCTGACCACTGAAT 180

DB 264 GGGGATTTGACTGTACTTCAATTTTGGACAAATCAGGAAGTGTCTGACCACTGAAT 323
 QY 181 GAAATCTAATCTTGTGGAACAGTTGGCTCACAAATTCATCAGCCACAGTTGAAGATG 240
 DB 324 GAAATCTAATCTTGTGGAACAGTTGGCTCACAAATTCATCAGCCACAGTTGAAGATG 383
 QY 241 TCCTTTATTTGTTTCTCCACCCGAGGAACAACCTTAAATGAATGACAGAAACAGAA 300
 DB 384 TCCTTTATTTGTTTCTCCACCCGAGGAACAACCTTAAATGAATGACAGAAACAGAA 443
 QY 301 CAATTCGTTAAAGCCCTTAAGAAATCTCAGAAAGTTCTGCAAGAGAGACACTTATG 360
 DB 444 CAATTCGTTAAAGCCCTTAAGAAATCTCAGAAAGTTCTGCAAGAGAGACACTTATG 503
 QY 361 CATGAAGATTTTGAAGGGCCAGTGAAGATTTATATGAAACAGACAAAGGTTACAGG 420
 DB 504 CATGAAGATTTTGAAGGGCCAGTGAAGATTTATATGAAACAGACAAAGGTTACAGG 563
 QY 421 ACAGCCAGGCTCATCTTGTGACTGATGAGAACTCCATGAAGATCTTTTCTAT 480
 DB 564 ACAGCCAGGCTCATCTTGTGACTGATGAGAACTCCATGAAGATCTTTTCTAT 623
 QY 481 TCAAGAGGAGGCTAATAGGTCTCGAGATCTTGTGCAATTTGTTACTGTGTGTG 540
 DB 624 TCAAGAGGAGGCTAATAGGTCTCGAGATCTTGTGCAATTTGTTACTGTGTGTG 683
 QY 541 AAAGATTTCAATGAGACACAGCTGGCCGGAATGGCGGACGTAAGATCATGTTTCCC 600
 DB 684 AAAGATTTCAATGAGACACAGCTGGCCGGAATGGCGGACGTAAGATCATGTTTCCC 743
 QY 601 GTGAATGACGGCTTTCAGGCTCTGCAAGGATCATCACTCAATTTTGAAGAGTCTG 660
 DB 744 GTGAATGACGGCTTTCAGGCTCTGCAAGGATCATCACTCAATTTTGAAGAGTCTG 803
 QY 661 ATCGAAATTTAGAGCTGAAACCATTCACATATGTGAGAGAGTCAATTTCAAGTTGTC 720
 DB 804 ATCGAAATTTAGAGCTGAAACCATTCACATATGTGAGAGAGTCAATTTCAAGTTGTC 863
 QY 721 GTGAGAGAAACGGCTTCCGACATGCGGCAAGGTGACAGGGCTCTGCAAGTTCAAG 780
 DB 864 GTGAGAGAAACGGCTTCCGACATGCGGCAAGGTGACAGGGCTCTGCAAGTTCAAG 923
 QY 781 ATCAATGACTCGGTCACTCAATGAGAAAGCCCTTTCTGTGAGAACACTTATTTACTG 840
 DB 924 ATCAATGACTCGGTCACTCAATGAGAAAGCCCTTTCTGTGAGAACACTTATTTACTG 983
 QY 841 TGTCCAGCGCTTATCTTAAAGAAAGTTGGCATGAAAGTTCGACACTCCAGTTCAGATGAC 900
 DB 984 TGTCCAGCGCTTATCTTAAAGAAAGTTGGCATGAAAGTTCGACACTCCAGTTCAGATGAC 1043
 QY 901 GATGGCCCTCTTTATCTCAAGTTCTGATATCAATCAACACACACTGTTCTGAAGGT 960
 DB 1044 GATGGCCCTCTTTATCTCAAGTTCTGATATCAATCAACACACACTGTTCTGAAGGT 1103
 QY 961 TCCATCTGCGCAATCGCCCTGCTGATCTGTTCTGCTGAGCCCTGCTCTCTG 1020
 DB 1104 TCCATCTGCGCAATCGCCCTGCTGATCTGTTCTGCTGAGCCCTGCTCTCTG 1163
 QY 1021 TGGTTCTGCGCCCTCTGCTGCACTGTGATTATCAAGAGGTCCCTCCACCCCTGCGAG 1080
 DB 1164 TGGTTCTGCGCCCTCTGCTGCACTGTGATTATCAAGAGGTCCCTCCACCCCTGCGAG 1223
 QY 1081 GAGAGTGAGGAAATTAATTAATAA 1094
 DB 1224 GAGAGTGAGGAAATTAATTAATAA 1237

RESULT 4

US-09-918-715-231
 ; Sequence 231, Application US/09918715
 ; Publication No. US20030017157A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Brad St. Croix

Query Match	98.8%	Score 1090.8	DB 10	Length 5540
Best Local Similarity	99.8%	Pred. No. 0		
Matches 1092	Conservative	0	Mismatches 2	Indels 0
			Gaps	0

QY	Db	Sequence	Score	DB	Length	Matches	Mismatches	Indels	Gaps
QY	Db	864 GTGAGAGGAAACGGCTTCCGACATGCCCCGACAGTGGACAGGGGTCTCTGACGCTTCAG	98.8%	1090.8	5540	2	0	0	0
QY	Db	781 ATCAATGACTCGGTCACACTCATATGAGAACCCCTTTCTGTGAAAGACATTATTTTACTG	99.8%	1090.8	5540	2	0	0	0
QY	Db	924 ATCAATGACTCGGTCACACTCATATGAGAACCCCTTTCTGTGAAAGTACTTATTTACTG	99.8%	1090.8	5540	2	0	0	0
QY	Db	841 TGTCCAGCGGCTTATCTTAAAGAAGTGGGATGAGAAAGCTCAGTCCAGGTCAGCATGAAC	99.8%	1090.8	5540	2	0	0	0
QY	Db	984 TGTCCAGCGGCTTATCTTAAAGAAGTGGGATGAGAAAGCTCAGTCCAGGTCAGCATGAAC	99.8%	1090.8	5540	2	0	0	0
QY	Db	901 GATGCGCTCTCTTTTATCTTCAGTTCGTGATCATCAACACACACTGTTCTGACGGT	99.8%	1090.8	5540	2	0	0	0
QY	Db	1044 GATGCGCTCTCTTTTATCTTCAGTTCGTGATCATCAACACACACTGTTCTGACGGT	99.8%	1090.8	5540	2	0	0	0
QY	Db	961 TCCATCTCGGCAATCGCCCTGCTGATTCCTGTCTGCTCTAGCCCTGCTCTCTCTG	99.8%	1090.8	5540	2	0	0	0
QY	Db	1104 TCCATCTCGGCAATCGCCCTGCTGATTCCTGTCTGCTCTAGCCCTGCTCTCTCTG	99.8%	1090.8	5540	2	0	0	0
QY	Db	1021 TGGTTCGAGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	99.8%	1090.8	5540	2	0	0	0
QY	Db	1164 TGGTTCGAGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	99.8%	1090.8	5540	2	0	0	0
QY	Db	1081 GAGAGTGAGGAAAA 1094	99.8%	1090.8	5540	2	0	0	0
QY	Db	1224 GAGAGTGAGGAAAA 1237	99.8%	1090.8	5540	2	0	0	0

```

Db      204 CTGGTGTCTACATCTGCGCGCGCAAGGGGAGGCGAGGAGGATGGGGGTCTCCAGCCTGTCTAC 263
Qy      121 GGGCGATTTTGAACCTGACTTCAATTTTGAACAATCAGAAAGTGTCTGACCACTGTGAAT 180
Db      264 GGGGGAATTTGACCTGTACTTCTTATTTTGGACAAATCAGAAAGTGTCTGACCACTGTGAAT 323
Qy      181 GAAATCTATTACTTTGTGGAACAGTTGGCTCAAAATTCATCAGCCCAAGTTGAGAATG 240
Db      324 GAAATCTATTACTTTGTGGAACAGTTGGCTCAAAATTCATCAGCCCAAGTTGAGAATG 383
Qy      241 TCCCTTATTTGTTTCTCAACCCGAGAAACAACCTTAATGAACCTGACAGAAAGAGAA 300
Db      384 TCCCTTATTTGTTTCTCAACCCGAGAAACAACCTTAATGAACCTGACAGAAAGAGAA 443
Qy      301 CAAATCCGTCAAGGCTAGAAAGAACTCAGAAAGTCTGCGCAGAGAGACACTTAATG 360
Db      444 CAAATCCGTCAAGGCTAGAAAGAACTCAGAAAGTCTGCGCAGAGAGACACTTAATG 503
Qy      361 CATGAAGATTTGAAGGGCCAGTGAAGATTTATTAAGAAAACAGAAAGGTACAGG 420
Db      504 CATGAAGATTTGAAGGGCCAGTGAAGATTTATTAAGAAAACAGAAAGGTACAGG 563
Qy      421 ACAGCCAGCGCTCATATTGCTTTGACTGATGAGAACTCCATGAAGATCTCTTTCTAT 480
Db      564 ACAGCCAGCGCTCATATTGCTTTGACTGATGAGAACTCCATGAAGATCTCTTTCTAT 623
Qy      481 TCAGAGAGGAGGCTATAGGCTCGAGATCTTGAGCAATGTTTCTGCTGTGGGTG 540
Db      624 TCAGAGAGGAGGCTATAGGCTCGAGATCTTGAGCAATGTTTCTGCTGTGGGTG 683
Qy      541 AAAGATTTCAATGAGACACAGCTGCGCGAATGCGGACAGTAAGATCATGTTTCCC 600
Db      684 AAAGATTTCAATGAGACACAGCTGCGCGAATGCGGACAGTAAGATCATGTTTCCC 743
Qy      601 GTGAATGAGGCTTTCAAGGCTCTGCAAGGATCATCACTCAATTTTGAAGAAAGTCTGC 660
Db      744 GTGAATGAGGCTTTCAAGGCTCTGCAAGGATCATCACTCAATTTTGAAGAAAGTCTGC 803
Qy      661 ATCGAAATTTAGAGAGTGAACCATCAACCAATATGTCAGAGAGATTTCAAGTTGTC 720
Db      804 ATCGAAATTTAGAGAGTGAACCATCAACCAATATGTCAGAGAGATTTCAAGTTGTC 863
Qy      721 GTGAGAGAAACGGCTTCCGACATGCGCGCAACGTGACAGGCTCTCTGCAAGTTCAAG 780
Db      864 GTGAGAGAAACGGCTTCCGACATGCGCGCAACGTGACAGGCTCTCTGCAAGTTCAAG 923
Qy      781 ATCAATGACTCGGTCACATCAATGAGAAAGCCTTTTCTGTGAGAAACACTTAATTAATG 840
Db      924 ATCAATGACTCGGTCACATCAATGAGAAAGCCTTTTCTGTGAGAAATTAATTAATG 983
Qy      841 TGTCCAGGGGCTATCTTAAAGAAAGTTGGATGAAGACTGCACTCCAGGTCAGACATGAAAC 900
Db      984 TGTCCAGGGGCTATCTTAAAGAAAGTTGGATGAAGACTGCACTCCAGGTCAGACATGAAAC 1043
Qy      901 GATGCGCTCTTTTATCTCAAGTTCTGTATCATCAACCAACACTGTTCTGACGGT 960
Db      1044 GATGCGCTCTTTTATCTCAAGTTCTGTATCATCAACCAACACTGTTCTGACGGT 1103
Qy      961 TCATCTCTGAGCATGCGCTGTGTGATCTGTCTCTGTAGCCCTGCTCTCTG 1020
Db      1104 TCATCTCTGAGCATGCGCTGTGTGATCTGTCTCTGTAGCCCTGCTCTCTG 1163
Qy      1021 TGGTTTCTGAGGCTCTGCTGCTGACATGTATATCAAGAGGTCTCTCAACCCCTGCGGAG 1080
Db      1164 TGGTTTCTGAGGCTCTGCTGCTGACATGTATATCAAGAGGTCTCTCAACCCCTGCGGAG 1223
Qy      1081 GAGAGTGAGGAAAA 1094
Db      1224 GAGAGTGAGGAAAA 1237

```

RESULT 6

```

US-10-474-794-176
; Sequence 176, Application US/10474794
; Publication No. US20040213793A1
; GENERAL INFORMATION:
; APPLICANT: Carson-Walter, Eleanor
; APPLICANT: St. Croix, Brad
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS
; FILE REFERENCE: 1107, 00179
; CURRENT APPLICATION NUMBER: US/10/474,794
; PRIORITY FILING DATE: 2003-10-14
; PRIOR APPLICATION NUMBER: 60/282,850
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 60/308,829
; PRIOR FILING DATE: 2001-08-01
; NUMBER OF SEQ. ID NOS: 359
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ. ID NO 176
; LENGTH: 5540
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-474-794-176

Query Match      98.8%; Score 1090.8; DB 20; Length 5540;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1092; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1  ATGGCCACGGCGGAGGCGGAGAGCCCTTGCGCATGCGCTTCCAGTGTCTCTTTGGCCACT 60
Db      144  ATGGCCACGGCGGAGGCGGAGAGCCCTTGCGCATGCGCTTCCAGTGTCTCTTTGGCCACT 203
Qy      61  CTGGTGTCTACATCTGCGCGCGCAAGGGGAGGCGAGGAGATGGGGGTCTCCAGCTGTCTAC 120
Db      204  CTGGTGTCTACATCTGCGCGCGCAAGGGGAGGCGAGGAGATGGGGGTCTCCAGCTGTCTAC 263
Qy      121  GGGGGAATTTGACCTGTACTTCTTATTTTGGACAAATCAGAAAGTGTCTGACCACTGTGAAT 180
Db      264  GGGGGAATTTGACCTGTACTTCTTATTTTGGACAAATCAGAAAGTGTCTGACCACTGTGAAT 323
Qy      181  GAAATCTATTACTTTGTGGAACAGTTGGCTCAAAATTCATCAGCCCAAGTTGAGAATG 240
Db      324  GAAATCTATTACTTTGTGGAACAGTTGGCTCAAAATTCATCAGCCCAAGTTGAGAATG 383
Qy      241  TCCCTTATTTGTTTCTCAACCCGAGAAACAACCTTAATGAACCTGACAGAAAGAGAA 300
Db      384  TCCCTTATTTGTTTCTCAACCCGAGAAACAACCTTAATGAACCTGACAGAAAGAGAA 443
Qy      301  CAAATCCGTCAAGGCTAGAAAGAACTCAGAAAGTCTGCGCAGAGAGACACTTAATG 360
Db      444  CAAATCCGTCAAGGCTAGAAAGAACTCAGAAAGTCTGCGCAGAGAGACACTTAATG 503
Qy      361  CATGAAGATTTGAAGGGCCAGTGAAGATTTATTAAGAAAACAGAAAGGTACAGG 420
Db      504  CATGAAGATTTGAAGGGCCAGTGAAGATTTATTAAGAAAACAGAAAGGTACAGG 563
Qy      421  ACAGCCAGGCTCATATTGCTTTGACTGATGAGAACTCCATGAAGATCTCTTTCTAT 480
Db      564  ACAGCCAGGCTCATATTGCTTTGACTGATGAGAACTCCATGAAGATCTCTTTCTAT 623
Qy      481  TCAGAGAGGAGGCTATAGGCTCGAGATCTTGAGCAATGTTTACTGTGTGGGTG 540
Db      624  TCAGAGAGGAGGCTATAGGCTCGAGATCTTGAGCAATGTTTACTGTGTGGGTG 683
Qy      541  AAAGATTTCAATGAGACACAGCTGCGCGAATGCGGACAGTAAGATCATGTTTCCC 600
Db      684  AAAGATTTCAATGAGACACAGCTGCGCGAATGCGGACAGTAAGATCATGTTTCCC 743
Qy      601  GTGAATGAGGCTTTGAGGCTCTGCAAGGATCATCACTCAATTTTGAAGAAAGTCTGC 660
Db      744  GTGAATGAGGCTTTGAGGCTCTGCAAGGATCATCACTCAATTTTGAAGAAAGTCTGC 803
Qy      661  ATCGAAATTTAGAGAGTGAACCATCAACCAATATGTCAGAGAGATTTCAAGTTGTC 720

```

```
Db 804 ATGGAATTTAGACAGTGAACCATCCATATGTGACAGAGAGTCAATTTCAAGTTGTC 863
Qy 721 GTGAGAGGAAACGGCTTCCGACATAGCCCGAAGGTGCGTCTGAGACTTCAAG 780
Db 864 GTGAGAGGAAACGGCTTCCGACATAGCCCGAAGGTGCGTCTGAGACTTCAAG 923
Qy 781 ATCAATGACTCGGTCACTCACTCAATGAGAAACCCCTTTCTGTGAGAGCACTTATTACTG 840
Db 924 ATCAATGACTCGGTCACTCACTCAATGAGAAACCCCTTTCTGTGAGAGTATTATTACTG 983
Qy 841 TGTCCAGCGCTTATCTTAAAGAGTTGGCATGAAAGCTGCACTCCAGTCAAGATGAAAC 900
Db 984 TGTCCAGCGCTTATCTTAAAGAGTTGGCATGAAAGCTGCACTCCAGTCAAGATGAAAC 1043
Qy 901 GATGGCTCTCTTTATCTCAGTCTGTGATCATACCAACAACATGTTGACGGT 960
Db 1044 GATGGCTCTCTTTATCTCAGTCTGTGATCATACCAACAACATGTTGACGGT 1103
Qy 961 TCCATCCTGAGCAATGCGCTGTGATCTGTTCTGCTAGCCCTGAGCTCTCTGAG 1020
Db 1104 TCCATCCTGAGCAATGCGCTGTGATCTGTTCTGCTAGCCCTGAGCTCTCTGAG 1163
Qy 1021 TGGTTCTGCGCCCTCTGCTGCACTGTGATTATCAAGAGGTCCCTCCACCCCTGCCGAG 1080
Db 1164 TGGTTCTGCGCCCTCTGCTGCACTGTGATTATCAAGAGGTCCCTCCACCCCTGCCGAG 1223
Qy 1081 GAGAGTGAGGAAA 1094
Db 1224 GAGAGTGAGGAAA 1237
```

RESULT 7
US-10-474-794-231

```
/ Sequence 231, Application US/10474794
/ Publication No. US20040213793A1
/ GENERAL INFORMATION:
/ APPLICANT: Carson-Walter, Eleanor
/ APPLICANT: St. Croix, Brad
/ APPLICANT: Vogelstein, Bert
/ APPLICANT: Kinzler, Kenneth
/ TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS
/ FILE REFERENCE: 1107, 00179
/ CURRENT APPLICATION NUMBER: US/10/474, 794
/ CURRENT FILING DATE: 2003-10-14
/ PRIOR APPLICATION NUMBER: 60/282, 850
/ PRIOR FILING DATE: 2001-04-11
/ PRIOR APPLICATION NUMBER: 60/308, 829
/ PRIOR FILING DATE: 2001-08-01
/ NUMBER OF SEQ ID NOS: 359
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 231
/ LENGTH: 5540
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ US-10-474-794-231
```

```
Query Match 98.8%; Score 1090.8; DB 20; Length 5540;
Beat Local Similarity 99.8%; Pred. No. 0;
Matches 1092; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 1 ATGGCCACGGCGGAGCGAGAGCCCTCGCATCGGCTTCCAGTGTCTCTTTGGCCACT 60
Db 144 ATGGCCACGGCGGAGCGAGAGCCCTCGCATCGGCTTCCAGTGTCTCTTTGGCCACT 203
Qy 61 CTGGTGTCTCATCTGCGCGGCGAAGGGGAGCGCGAGAGATGGGGGTCCAGCTCTAC 120
Db 204 CTGGTGTCTCATCTGCGCGGCGAAGGGGAGCGCGAGAGATGGGGGTCCAGCTCTAC 263
Qy 121 GGGCGATTTGACCTGTACTTCAATTTTGGACAAATCAGGAAGTGTGTGACCACTGGAAAT 180
Db 264 GGGCGATTTGACCTGTACTTCAATTTTGGACAAATCAGGAAGTGTGTGACCACTGGAAAT 323
```

```
Qy 181 GAAATCTATTACTTTGTGAAACAGTTGGCTCACAAAATTCATACGCCACAGTTGAGAAATG 240
Db 324 GAAATCTATTACTTTGTGAAACAGTTGGCTCACAAAATTCATACGCCACAGTTGAGAAATG 383
Qy 241 TCCCTTATTTGTTTCTCCACCGGAGGAAACACTTAAATGAACTGACAGAGACAGAGAA 300
Db 384 TCCCTTATTTGTTTCTCCACCGGAGGAAACACTTAAATGAACTGACAGAGACAGAGAA 443
Qy 301 CAAATCCGTCAAGGCTAGAGAAATCCAGAAAGTTGTGCGCAGAGAGACACTTACATG 360
Db 444 CAAATCCGTCAAGGCTAGAGAAATCCAGAAAGTTGTGCGCAGAGAGACACTTACATG 503
Qy 361 CATGAAAGATTTGAAAGGGCCAGTGACAGATTTATTGAAAGACAGACAGGGTACAG 420
Db 504 CATGAAAGATTTGAAAGGGCCAGTGACAGATTTATTGAAAGACAGACAGGGTACAG 563
Qy 421 ACAGCGAGGTATCATGCTGCTTGACTGATGAGAACTCATGAGATCTCTTTTCTAT 480
Db 564 ACAGCGAGGTATCATGCTTGTGACTGATGAGAACTCATGAGATCTCTTTTCTAT 623
Qy 481 TCAGAGAGGAGGCTTAATAGTCTGAGATCTTGTGCAATGTTTACTGTGTGTGTG 540
Db 624 TCAGAGAGGAGGCTTAATAGTCTGAGATCTTGTGTGCAATGTTTACTGTGTGTGTG 583
Qy 541 AAAGATTTCAATGAGACACAGCTGGCCGGAATTTGGCGAAGTAAATGATGATGTTTCCC 600
Db 684 AAAGATTTCAATGAGACACAGCTGGCCGGAATTTGGCGAAGTAAATGATGATGTTTCCC 743
Qy 601 GTGAATGACGGCTTTCAGGCTGCGAAGGATCATCTCAATTTTGAAGAGTCTCTG 660
Db 744 GTGAATGACGGCTTTCAGGCTGCGAAGGATCATCTCAATTTTGAAGAGTCTCTG 803
Qy 661 ATCGAAATTTCTAGCAGCTGAAACATCCACATATGTGACAGAGAGTCAATTCAGTTGTC 720
Db 804 ATCGAAATTTCTAGCAGCTGAAACATCCATATGTGACAGAGAGTCAATTCAGTTGTC 863
Qy 721 GTGAGAGGAAACGGCTTCCGACATGCGCCGCAACGTGGAACAGGTCCTCTGAGCTTCAAG 780
Db 864 GTGAGAGGAAACGGCTTCCGACATGCGCCGCAACGTGGAACAGGTCCTCTGAGCTTCAAG 923
Qy 781 ATCAATGACTCGGTCACTCACTCAATGAGAAACCCCTTTCTGTGAGAGCACTTATTACTG 840
Db 924 ATCAATGACTCGGTCACTCACTCAATGAGAAACCCCTTTCTGTGAGAGTATTATTACTG 983
Qy 901 GATGGCTCTCTTTATCTCAGTCTGTGATCATACCAACAACACACTGTTGTACGGT 960
Db 1044 GATGGCTCTCTTTATCTCAGTCTGTGATCATACCAACAACACACTGTTGTACGGT 1103
Qy 961 TCCATCCTGAGCAATGCGCTGTGATCTGTTCTGCTAGCCCTGAGCTCTCTGAG 1020
Db 1104 TCCATCCTGAGCAATGCGCTGTGATCTGTTCTGCTAGCCCTGAGCTCTCTGAG 1163
Qy 1021 TGGTTCTGCGCCCTCTGCTGCACTGTGATTATCAAGAGGTCCCTCCACCCCTGCCGAG 1080
Db 1164 TGGTTCTGCGCCCTCTGCTGCACTGTGATTATCAAGAGGTCCCTCCACCCCTGCCGAG 1223
Qy 1081 GAGAGTGAGGAAA 1094
Db 1224 GAGAGTGAGGAAA 1237
```

RESULT 8
US-10-038-307-17

```
/ Sequence 17, Application US/10038307
/ Publication No. US20030134786A1
/ GENERAL INFORMATION:
/ APPLICANT: James B. ROTHMAN
/ APPLICANT: Theresa L. O'KEEFE
/ APPLICANT: Engin OZKANNAK
```


? APPLICANT: Judith J. HEALEY
 ? TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
 ? FILE REFERENCE: 7853-253-999
 ? CURRENT APPLICATION NUMBER: US/10/038.307
 ? CURRENT FILING DATE: 2002-06-28
 ? NUMBER OF SEQ ID NOS: 26
 ? SOFTWARE: PasteSeq for Windows Version 4.0
 ? SEQ ID NO 17
 ? LENGTH: 1674
 ? TYPE: DNA
 ? ORGANISM: Homo sapiens
 ? US-10-038-307-17

Query Match	87.0%;	Score 960.8;	DB 15;	Length 1674;
Best Local Similarity	99.8%;	Pred. No. 3.9e-303;		
Matches 962; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0

QY	ATGGCAAGGGGAGAGGAGAGGCGCTCGGCATCGGCTTCAGATGGCTCTCTTTGGCCACT	60
Db	ATGGCCAGGGGAGAGGAGAGGCGCTCGGCATCGGCTTCAGATGGCTCTCTTTGGCCACT	72
QY	CTGGTGCTCATCTGGCCCGGCGAAGGGGAGCGACGAGAGAGATGGGGGTTCAGCTGTCTAC	120
Db	CTGGTGCTCATCTGGCGCGGCGAAGGGGAGCGACGAGAGAGATGGGGGTTCAGCTGTCTAC	132
QY	GGCGGATTTGACCTGTACTTCATTTTGGACAATCAGGAAGTGTGCTGCAACCTGGAAAT	180
Db	GGCGGATTTGACCTGTACTTCATTTTGGACAATCAGGAAGTGTGCTGCAACCTGGAAAT	192
QY	GAATTCATTACTTGTGGAAACAGTGGCTCAAAATTCATGAGCCACAGTTGGAATG	240
Db	GAATTCATTACTTGTGGAAACAGTGGCTCAAAATTCATGAGCCACAGTTGGAATG	252
QY	TCCTTTATGTTTTCTCCAGCCGAGAAACAACCTTAATGAACTGACAGAAACAGAGA	300
Db	TCCTTTATGTTTTCTCCAGCCGAGAAACAACCTTAATGAACTGACAGAAACAGAGA	312
QY	CAATTCGCTCAAGGCTTAGAAGAACTCCAGAAAGTTCTGCCAGAGAGACACTTAACATG	360
Db	CAATTCGCTCAAGGCTTAGAAGAACTCCAGAAAGTTCTGCCAGAGAGACACTTAACATG	372
QY	CATGAAGATTGGAAAGGGCCAGTGAAGAGATTTATTAAGAAAAGACAAGGGGTACAGG	420
Db	CATGAAGATTGGAAAGGGCCAGTGAAGAGATTTATTAAGAAAAGACAAGGGGTACAGG	432
QY	ACAGCCAGCGTCATCATTTGCTTTGACTGATGAGAGACTCAATGAAGATCTCTTTTCTAT	480
Db	ACAGCCAGCGTCATCATTTGCTTTGACTGATGAGAGACTCAATGAAGATCTCTTTTCTAT	492
QY	TCAGAGAGGGAGGCTAATAGGTTCTCGAATCTTGGTGCAATTTGTTACTGTGTGGTGTG	540
Db	TCAGAGAGGGAGGCTAATAGGTTCTCGAATCTTGGTGCAATTTGTTACTGTGTGGTGTG	552
QY	AAAGATTTCAATGAGACACAGCTGGCCCGGATTTGGGACAGTAAAGATCATGTGTTCCC	600
Db	AAAGATTTCAATGAGACACAGCTGGCCCGGATTTGGGACAGTAAAGATCATGTGTTCCC	612
QY	GTEGAATGACGGCTTCAGGCTCTGCAAGGCATCATCACTCAATTTGAAGAACTCTGC	660
Db	GTEGAATGACGGCTTCAGGCTCTGCAAGGCATCATCACTCAATTTGAAGAACTCTGC	672
QY	ATCGAAATTTTCAGAGAGCTGAACATCCACCAATATGTGAGAGAGATCATTTCAAGTTGTC	720
Db	ATCGAAATTTTCAGAGAGCTGAACATCCACCAATATGTGAGAGAGATCATTTCAAGTTGTC	732
QY	GTCAGAGGAAACGGCTTCCGACATGCCCGCAACGTGACAGGGGTCTCTGACGCTTCAAG	780
Db	GTCAGAGGAAACGGCTTCCGACATGCCCGCAACGTGACAGGGGTCTCTGACGCTTCAAG	792
QY	ATCAATAGCTCGGTCAACATCAAGAAAGCCCTTTTCTGTGGAAAGACATTAATTACTG	840
Db	ATCAATAGCTCGGTCAACATCAAGAAAGCCCTTTTCTGTGGAAAGATTAATTACTG	852

Qy	841	TCGCAAGCCGCTATCTTTAAAGAATTGGCATGAAAGCTGCATCCGAGTCAGCATGAAC	900
Db	853	TGTCGACGCCCTATCTTTAAAGAATTGGCATGAAAGCTGCATCCGAGTCAGCATGAAC	912
Qy	901	GATGGAGCTCTTTTATCTCGAGTCTGTGCATCATCAACACACACACTGTTCTGACGCT	960
Db	913	GATGGAGCTCTTTTATCTCGAGTCTGTGCATCATCAACACACACACTGTTCTGACGCT	972
Qy	961	TCGA	964
Db	973	CCCA	976

RESULT 9
US-10-201-292-17
Sequence 17 Application file/10201292

Query	Match	Best Local Similarity	Matches	Score	Pred. No.	DB	Length	Indels	Gaps
GENERAL INFORMATION:									
APPLICANT: James B. ROTTMAN									
APPLICANT: Theresa L. O'KEEFE									
APPLICANT: Engin OZKAYNAK									
APPLICANT: Judith J. HEALEY									
TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods									
FILE REFERENCE: 7853-253-999									
CURRENT APPLICATION NUMBER: US/10/201,292									
CURRENT FILING DATE: 2003-02-14									
NUMBER OF SEQ ID NOS: 36									
SOFTWARE: FastSeq for Windows Version 4.0									
SEQ ID NO 17									
LENGTH: 1674									
TYPE: DNA									
ORGANISM: Homo sapiens									
US-10-201-292-17									
Query Match	87.0%	Score 960.8;	DB 15;	Length 1674;					
Best Local Similarity	99.8%;	Pred. No. 3,9e-303;							
Matches 962;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;					
1	ATGGCCACGGCGGAGCGGAGAGCCCTGGCGATCGGCTTCAGTGGCTCTCTTGGCCACT	60							
13	ATGGCCACGGGAGCGGAGAGCCCTGGCGATCGGCTTCAGTGGCTCTCTTGGCCACT	72							
61	CTGGTGTCTCATCTGCGCGCGGCAAGGGGGAGCGAGGAGAGATGGGGGCTCCAGCTGTCTAC	120							
73	CTGGTGTCTCATCTGCGCGCGGCAAGGGGGAGCGAGGAGAGATGGGGGCTCCAGCTGTCTAC	132							
121	GGCGGATTTGACCTGTACTTCTTGTGGACAAATCAGGAAGTGTGTGACCACTGGAAAT	180							
133	GGCGGATTTGACCTGTACTTCTTGTGGACAAATCAGGAAGTGTGTGACCACTGGAAAT	192							
181	GAATCTATTACTTTTGGAGAACAGTTGGCTCACAAATTCATCAGCCCAAGTTGGAAATG	240							
193	GAATCTATTACTTTTGGAGAACAGTTGGCTCACAAATTCATCAGCCCAAGTTGGAAATG	252							
241	TCTCTTATGTGTTTCTCCACCAGGAAACAACCTTAATGAATGACAGAAACAGAGA	300							
253	TCTCTTATGTGTTTCTCCACCAGGAAACAACCTTAATGAATGACAGAAACAGAGA	312							
301	CAATCGGTCAAGGCTTAGAAGAACTCCAGAAAGTTCGACAGAGGAGACACTTACATG	360							
313	CAATCGGTCAAGGCTTAGAAGAACTCCAGAAAGTTCGACAGAGGAGACACTTACATG	372							
361	CATGAAGATTTGAAAGGGCCAGTAGACAGATTTATTATGAAAACAGCAAGGGTACAGG	420							
373	CATGAAGATTTGAAAGGGCCAGTAGACAGATTTATTATGAAAACAGCAAGGGTACAGG	432							
421	ACAGCCAGCGCATCATTTGCTTTGACGTAGAGAGAACTCCATGAAGATCTCTTTTCAT	480							
433	ACAGCCAGCGCATCATTTGCTTTGACGTAGAGAGAACTCCATGAAGATCTCTTTTCAT	492							
481	TCAGAGAGGAGGCTAATAGGCTCTAGAGATCTTGATGCAATGTTTATCTGTGTGGTGTG	540							
493	TCAGAGAGGAGGCTAATAGGCTCTAGAGATCTTGATGCAATGTTTATCTGTGTGGTGTG	552							

QY 541 AAAGATTTCATGAGACAGAGCTGGCCCGGATTCGCGACAGTAAGATCATGTGTTCC 600
DB 553 AAAGATTTCATGAGACAGAGCTGGCCCGGATTCGCGACAGTAAGATCATGTGTTCC 612
QY 601 GTGATGAGCGCTTTCAGGCTCTGCGAAGCATCATCAATTTTGAAGAAAGTCTGC 660
DB 613 GTGATGAGCGCTTTCAGGCTCTGCGAAGCATCATCAATTTTGAAGAAAGTCTGC 672
QY 661 ATGGAATTCAGAGCTGAAACCATCCACATATGTCAGAGAGATCATTTCAAGTTGTC 720
DB 673 ATGGAATTCAGAGCTGAAACCATCCACATATGTCAGAGAGATCATTTCAAGTTGTC 732
QY 721 GTGAGAGAAACGGCTTCCGACATGCGCGCAACGTGACAGGGTCTCTGCAAGTTCAAG 780
DB 733 GTGAGAGAAACGGCTTCCGACATGCGCGCAACGTGACAGGGTCTCTGCAAGTTCAAG 792
QY 781 ATCAATGACTCGGTCACTCACTCAATGAGAAAGCCCTTTCTGTGAAAGACATTTATTA 840
DB 793 ATCAATGACTCGGTCACTCACTCAATGAGAAAGCCCTTTCTGTGAAAGATTTATTA 852
QY 841 TGTCCAGCGCTATCTTAAAGAAAGTTGGCATGAAGCTGACCTCCAGTCAGCATGAC 900
DB 853 TGTCCAGCGCTATCTTAAAGAAAGTTGGCATGAAGCTGACCTCCAGTCAGCATGAC 912
QY 901 GATGGCTCTCTTTATCTCCAGTTCTGTCAATCAACACACACTGTTCTGACGCT 960
DB 913 GATGGCTCTCTTTATCTCCAGTTCTGTCAATCAACACACACTGTTCTGACGCT 972
QY 961 TCCTA 964
DB 973 CCCA 976

RESULT 10
US-10-038-307-13
; Sequence 13, Application US/10038307
; Publication No. US20030134786A1
; GENERAL INFORMATION:
; APPLICANT: James B. ROTTMAN
; APPLICANT: Theresa L. O'KEEFE
; APPLICANT: Engin OZKAYNAK
; APPLICANT: Judith J. HEALEY
; TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
; FILE REFERENCE: 7853-253-999
; CURRENT APPLICATION NUMBER: US/10/038,307
; CURRENT FILING DATE: 2002-06-28
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1650
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-307-13

Query Match 86.3%; Score 952.4; DB 15; Length 1650;
Best Local Similarity 99.4%; Pred. No. 2.2e-300;
Matches 956; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 ATGGCCACGGCGGAGAGAGCCCTGGCATGGCTTCCAGTGGCTCTCTTTGGCCACT 60
DB 1 ATGGCCACGGCGGAGAGAGCCCTGGCATGGCTTCCAGTGGCTCTCTTTGGCCACT 60
QY 61 CTGGTGTCTCATCTGGCGCGGAGAGAGGAGGAGATGGGGGTCCAGCTGCTAC 120
DB 61 CTGGTGTCTCATCTGGCGCGGAGAGAGGAGGAGATGGGGGTCCAGCTGCTAC 120
QY 121 GGGCGAATTTGACTGTACTTATTTTGGACAATCAGAACTGTGTGCACTGTGAAT 180
DB 121 GGGCGAATTTGACTGTACTTATTTTGGACAATCAGAACTGTGTGCACTGTGAAT 180
QY 181 GAAATCTATTACTTTGTGGAACAGTTGGCTCACAAAATTCATGAGCCCAAGTTGAATG 240
DB 181 GAAATCTATTACTTTGTGGAACAGTTGGCTCACAAAATTCATGAGCCCAAGTTGAATG 240

DB 181 GAAATCTATTACTTTGTGGAACAGTTGGCTCACAAAATTCATGAGCCCAAGTTGAATG 240
QY 241 TCCTTTATTTGTTTCTTCACCCGAGGAAACCTTAATGAACCTGACAGAGACAGAGAA 300
DB 241 TCCTTTATTTGTTTCTTCACCCGAGGAAACCTTAATGAACCTGACAGAGACAGAGAA 300
QY 301 CAAATCCGTCAAGGCTTAAGAGATCTCCAGAAAGTTCTGTCAGAGAGAGACATTAATG 360
DB 301 CAAATCCGTCAAGGCTTAAGAGATCTCCAGAAAGTTCTGTCAGAGAGAGACATTAATG 360
QY 361 CATGAAGATTTGAAGAGGCGCAGTGAAGATTTATTAAGAAAAGAGAGGTTGACAG 420
DB 361 CATGAAGATTTGAAGAGGCGCAGTGAAGATTTATTAAGAAAAGAGAGGTTGACAG 420
QY 421 ACAGCCAGCGTATCATTTGCTTTGACATGAGAACTCCATGAAAGATCTTTTCTAT 480
DB 421 ACAGCCAGCGTATCATTTGCTTTGACATGAGAACTCCATGAAAGATCTTTTCTAT 480
QY 481 TCAGAGAGGAGGCTTAATGAGTCTGAGATCTTTGTGCAATTTGTACTGTGTGTTG 540
DB 481 TCAGAGAGGAGGCTTAATGAGTCTGAGATCTTTGTGCAATTTGTACTGTGTGTTG 540
QY 541 AAAGATTTCATGAGACAGAGCTGGCCCGGATTCGCGACAGTAAGATCATGTGTTCC 600
DB 541 AAAGATTTCATGAGACAGAGCTGGCCCGGATTCGCGACAGTAAGATCATGTGTTCC 600
QY 601 GTGATGAGCGCTTTCAGGCTCTGCGAAGCATCATCAATTTTGAAGAAAGTCTGC 660
DB 601 GTGATGAGCGCTTTCAGGCTCTGCGAAGCATCATCAATTTTGAAGAAAGTCTGC 660
QY 661 ATGGAATTCAGAGCTGAAACCATCCACATATGTCAGAGAGATCATTTCAAGTTGTC 720
DB 661 ATGGAATTCAGAGCTGAAACCATCCACATATGTCAGAGAGATCATTTCAAGTTGTC 720
QY 721 GTGAGAGAAACGGCTTCCGACATGCGCGCAACGTGACAGGGTCTCTGCAAGTTCAAG 780
DB 721 GTGAGAGAAACGGCTTCCGACATGCGCGCAACGTGACAGGGTCTCTGCAAGTTCAAG 780
QY 781 ATCAATGACTCGGTCACTCACTCAATGAGAAAGCCCTTTCTGTGAAAGACTTTATTA 840
DB 781 ATCAATGACTCGGTCACTCACTCAATGAGAAAGCCCTTTCTGTGAAAGATTTATTA 840
QY 841 TGTCCAGCGCTATCTTAAAGAAAGTTGGCATGAAGCTGCACTCCAGTCAGCATGAC 900
DB 841 TGTCCAGCGCTATCTTAAAGAAAGTTGGCATGAAGCTGCACTCCAGTCAGCATGAC 900
QY 901 GATGGCTCTCTTTATCTCCAGTTCTGTCAATCAACACACACTGTTCTGACGCT 960
DB 901 GATGGCTCTCTTTATCTCCAGTTCTGTCAATCAACACACACTGTTCTGACGCT 960
QY 961 TC 962
DB 961 TC 962

RESULT 11
US-10-038-307-15
; Sequence 15, Application US/10038307
; Publication No. US20030134786A1
; GENERAL INFORMATION:
; APPLICANT: James B. ROTTMAN
; APPLICANT: Theresa L. O'KEEFE
; APPLICANT: Engin OZKAYNAK
; APPLICANT: Judith J. HEALEY
; TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
; FILE REFERENCE: 7853-253-999
; CURRENT APPLICATION NUMBER: US/10/038,307
; CURRENT FILING DATE: 2002-06-28
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 1650
; TYPE: DNA

! ORGANISM: Homo sapiens
US-10-038-307-15

Query Match 86.3%; Score 952.4; DB 15; Length 1650;
Best Local Similarity 99.4%; Pred. No. 2.2e-300;
Matches 956; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
QY 1 ATGGCCACGGCGGAGAGAGCCCTGGCATCGGCTTCCAGTGGCTCTCTTTGGCCACT 60
DB 1 ATGGCCACGGCGGAGAGAGCCCTGGCATCGGCTTCCAGTGGCTCTCTTTGGCCACT 60
QY 61 CTGGGCTCATCTGGCCGGGCAAGGGGAGCGCAGGAGATGGGGGTCAGCCCTGCTAC 120
DB 61 CTGGGCTCATCTGGCCGGGCAAGGGGAGCGCAGGAGATGGGGGTCAGCCCTGCTAC 120
QY 121 GCGGATTTGACCTGTACTCTTATTTGACAAATCAGGAAGTGTCTGCACCACTGGAAT 180
DB 121 GCGGATTTGACCTGTACTCTTATTTGACAAATCAGGAAGTGTCTGCACCACTGGAAT 180
QY 181 GAAATCTATTACTTTGTGGAACAGTTGGCTCAAAATTCATCAGCCCAAGTTGAAATG 240
DB 181 GAAATCTATTACTTTGTGGAACAGTTGGCTCAAAATTCATCAGCCCAAGTTGAAATG 240
QY 241 TCCCTTATGTTTTCTCACCAGGAAACAACCTTAATGAAACTGACAGAAACAGAGAA 300
DB 241 TCCCTTATGTTTTCTCACCAGGAAACAACCTTAATGAAACTGACAGAAACAGAGAA 300
QY 301 CAAATCCGTCAAGGCTTAGAAGAACTCCAGAAAGTTCTCCAGAGAGAGACACTTACATG 360
DB 301 CAAATCCGTCAAGGCTTAGAAGAACTCCAGAAAGTTCTCCAGAGAGAGACACTTACATG 360
QY 361 CATGAAGATTTGAAGGGCCAGTGAAGCAATTAATGAAACAGACAGGATACAG 420
DB 361 CATGAAGATTTGAAGGGCCAGTGAAGCAATTAATGAAACAGACAGGATACAG 420
QY 421 ACAGCCAGGCTCATCTTGTGACTGATGAGAACTCCATGAAATCTCTTTTCTAT 480
DB 421 ACAGCCAGGCTCATCTTGTGACTGATGAGAACTCCATGAAATCTCTTTTCTAT 480
QY 481 TCAGAGAGGAGGCTTAATAGGCTCGAGATCTGGTGCATTTGTTTACTGTGGTGTG 540
DB 481 TCAGAGAGGAGGCTTAATAGGCTCGAGATCTGGTGCATTTGTTTACTGTGGTGTG 540
QY 541 AAAGATTTCAATGAGACACAGCTGGCCGAGTTGGACAGTGAAGATCATGTGTTCCC 600
DB 541 AAAGATTTCAATGAGACACAGCTGGCCGAGTTGGACAGTGAAGATCATGTGTTCCC 600
QY 601 GTGAATGACGGCTTCAAGCTCTGACAGGATCATCACTCAATTTTGAAGAAGTCTGCG 660
DB 601 GTGAATGACGGCTTCAAGCTCTGACAGGATCATCACTCAATTTTGAAGAAGTCTGCG 660
QY 661 ATCGAAATTTCAAGAGCTGAACATCCACATATGTGAGAGAGTCACTTCAAGTTGTC 720
DB 661 ATCGAAATTTCAAGAGCTGAACATCCACATATGTGAGAGAGTCACTTCAAGTTGTC 720
QY 721 GTGAGAGGAAACGGCTTCCGACATGCCCCGCAACGTGACAGGCTCTCTGACGCTTCAAG 780
DB 721 GTGAGAGGAAACGGCTTCCGACATGCCCCGCAACGTGACAGGCTCTCTGACGCTTCAAG 780
QY 781 ATCAATGACTCGGTCACTCAATGAGAAAGCCCTTTCTGTGAGAGACACTTAATTTACTG 840
DB 781 ATCAATGACTCGGTCACTCAATGAGAAAGCCCTTTCTGTGAGAGACACTTAATTTACTG 840
QY 841 TGTCAAGGCTATCTTAAAGAGTTGGCATGAAAGTGTGACCTCCAGGTGAGATGAAAC 900
DB 841 TGTCAAGGCTATCTTAAAGAGTTGGCATGAAAGTGTGACCTCCAGGTGAGATGAAAC 900
QY 901 GATGGCTCTCTTTATCTCAGTTCTGTCAATCAACCACTGTTCTGACGGT 960
DB 901 GATGGCTCTCTTTATCTCAGTTCTGTCAATCAACCACTGTTCTTCCCAA 960
QY 961 TC 962
DB 961 TC 962
```

DB 961 TC 962

```
RESULT 12
US-10-292-13
; Sequence 13, Application US/10201292
; Publication No. US20030144193A1
; GENERAL INFORMATION:
; APPLICANT: James B. ROTHMAN
; APPLICANT: Theresa L. O'KEEFE
; APPLICANT: Engin OZKANAK
; APPLICANT: Judith J. HENLEY
; TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
; FILE REFERENCE: 7853-253-999
; CURRENT APPLICATION NUMBER: US/10/201,292
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1650
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-292-13
```

Query Match 86.3%; Score 952.4; DB 15; Length 1650;
Best Local Similarity 99.4%; Pred. No. 2.2e-300;
Matches 956; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
QY 1 ATGGCCACGGCGGAGAGAGCCCTGGCATCGGCTTCCAGTGGCTCTCTTTGGCCACT 60
DB 1 ATGGCCACGGCGGAGAGAGCCCTGGCATCGGCTTCCAGTGGCTCTCTTTGGCCACT 60
QY 61 CTGGGCTCATCTGGCCGGGCAAGGGGAGCGCAGGAGATGGGGGTCAGCCCTGCTAC 120
DB 61 CTGGGCTCATCTGGCCGGGCAAGGGGAGCGCAGGAGATGGGGGTCAGCCCTGCTAC 120
QY 121 GCGGATTTGACCTGTACTCTTATTTGACAAATCAGGAAGTGTCTGCACCACTGGAAT 180
DB 121 GCGGATTTGACCTGTACTCTTATTTGACAAATCAGGAAGTGTCTGCACCACTGGAAT 180
QY 181 GAAATCTATTACTTTGTGGAACAGTTGGCTCAAAATTCATCAGCCCAAGTTGAAATG 240
DB 181 GAAATCTATTACTTTGTGGAACAGTTGGCTCAAAATTCATCAGCCCAAGTTGAAATG 240
QY 241 TCCCTTATGTTTTCTCACCAGGAAACAACCTTAATGAAACTGACAGAAACAGAGAA 300
DB 241 TCCCTTATGTTTTCTCACCAGGAAACAACCTTAATGAAACTGACAGAAACAGAGAA 300
QY 301 CAAATCCGTCAAGGCTTAGAAGAACTCCAGAAAGTTCTCCAGAGAGAGACACTTACATG 360
DB 301 CAAATCCGTCAAGGCTTAGAAGAACTCCAGAAAGTTCTCCAGAGAGAGACACTTACATG 360
QY 361 CATGAAGATTTGAAGGGCCAGTGAAGCAATTAATGAAACAGACAGGATACAG 420
DB 361 CATGAAGATTTGAAGGGCCAGTGAAGCAATTAATGAAACAGACAGGATACAG 420
QY 421 ACAGCCAGGCTCATCTTGTGACTGATGAGAACTCCATGAAAGTCTCTTTTCTAT 480
DB 421 ACAGCCAGGCTCATCTTGTGACTGATGAGAACTCCATGAAAGTCTCTTTTCTAT 480
QY 481 TCAGAGAGGAGGCTTAATAGGCTCGAGATCTTGTGCAATTTGTTTACTGTGTGTG 540
DB 481 TCAGAGAGGAGGCTTAATAGGCTCGAGATCTTGTGCAATTTGTTTACTGTGTGTG 540
QY 541 AAAGATTTCAATGAGACACAGCTGGCCGAGTTGGACAGTGAAGATCATGTGTTCCC 600
DB 541 AAAGATTTCAATGAGACACAGCTGGCCGAGTTGGACAGTGAAGATCATGTGTTCCC 600
QY 601 GTGAATGACGGCTTCAAGCTCTGACAGGATCATCACTCAATTTTGAAGAAGTCTGCG 660
DB 601 GTGAATGACGGCTTCAAGCTCTGACAGGATCATCACTCAATTTTGAAGAAGTCTGCG 660
QY 661 ATCGAAATTTCAAGAGCTGAACATCCACATATGTGAGAGAGTCACTTCAAGTTGTC 720
DB 661 ATCGAAATTTCAAGAGCTGAACATCCACATATGTGAGAGAGTCACTTCAAGTTGTC 720
```



```
QY 61 CTGGTCTCATCTGCGCGGCAAGGGGACGACAGAGAGATGAGGGGCTCCAGCTGTAC 120
DB 61 CTGGTCTCATCTGCGCGGCAAGGGGACGACAGAGAGATGAGGGGCTCCAGCTGTAC 120
QY 121 GCGGATTTGACCTGTACTTCACTTTTGGACAAATCAGGAAGTGTCTGACCACTGGAAT 180
DB 121 GCGGATTTGACCTGTACTTCACTTTTGGACAAATCAGGAAGTGTCTGACCACTGGAAT 180
QY 181 GAAATCTATTACTTGTGGACAGTTGGCTCAAAATTCATCAGGCCACAGTTGAGAAATG 240
DB 181 GAAATCTATTACTTGTGGACAGTTGGCTCAAAATTCATCAGGCCACAGTTGAGAAATG 240
QY 241 TCCTTTATTTGTTTCTCCACCCGAGAAACACCTTAAATGAAATGAAATGAAATGAAATG 300
DB 241 TCCTTTATTTGTTTCTCCACCCGAGAAACACCTTAAATGAAATGAAATGAAATGAAATG 300
QY 301 CAATTCCTTCAAGGCTTGAAGAACTCCAGAAAGTTCTGCGAGAGAGACCTTACATG 360
DB 301 CAATTCCTTCAAGGCTTGAAGAACTCCAGAAAGTTCTGCGAGAGAGACCTTACATG 360
QY 361 CATGAAGATTTGAAGGGCCAGTGAAGATTTATTAATGAAACAGACAAGGGTACAG 420
DB 361 CATGAAGATTTGAAGGGCCAGTGAAGATTTATTAATGAAACAGACAAGGGTACAG 420
QY 421 ACAGCAGCGTCTCATCTGCTTGAATGAGAACTTCATGAAGATCTCTTTTCTAT 480
DB 421 ACAGCAGCGTCTCATCTGCTTGAATGAGAACTTCATGAAGATCTCTTTTCTAT 480
QY 481 TCAGAGAGAGAGGCTTAATAGGCTCTGAGATCTTGGTCAATTTGTTACTGTGTGTG 540
DB 481 TCAGAGAGAGAGGCTTAATAGGCTCTGAGATCTTGGTCAATTTGTTACTGTGTGTG 540
QY 541 AAAATTTCAATGAGACACAGCTGCGGATTTGGGACAGTAAGATCATGTGTTCC 600
DB 541 AAAATTTCAATGAGACACAGCTGCGGATTTGGGACAGTAAGATCATGTGTTCC 600
QY 601 GTGAATGACGGCTTCAAGCTCTGCAAGGATCATCACTCAATTTTGAAGAGTCTGCG 660
DB 601 GTGAATGACGGCTTCAAGCTCTGCAAGGATCATCACTCAATTTTGAAGAGTCTGCG 660
QY 661 ATCGAAATTTCTAGCAGCTGAAACATCCATATGTGCAAGAGTCACTTCAAGTGTG 720
DB 661 ATCGAAATTTCTAGCAGCTGAAACATCCATATGTGCAAGAGTCACTTCAAGTGTG 720
QY 721 GTGAGAGAAACGGCTTCCGACATGCGCGCAACGTGACAGGGTCTCTGCAAGTCTCAAG 780
DB 721 GTGAGAGAAACGGCTTCCGACATGCGCGCAACGTGACAGGGTCTCTGCAAGTCTCAAG 780
QY 781 ATCAATGACCTGCTGACACTCAATGAGAAAGCTTTCTGTGAGAAACATTAATTAATG 840
DB 781 ATCAATGACCTGCTGACACTCAATGAGAAAGCTTTCTGTGAGAAACATTAATTAATG 840
QY 841 TGTCCAGCGCTATCTTAAAGAGTGGCAATGAAAGCTGCACTCAGGTGAGAGTGAAC 900
DB 841 TGTCCAGCGCTATCTTAAAGAGTGGCAATGAAAGCTGCACTCAGGTGAGAGTGAAC 900
QY 901 GATGCGCTCTCTTTATCTCAAGTCTGTCATCACTCAACCAACACACTCTTCTCCAA 960
DB 901 GATGCGCTCTCTTTATCTCAAGTCTGTCATCACTCAACCAACACACTCTTCTCCAA 960
QY 961 TC 962
DB 961 TC 962
```

```
RESULT 15
US-10-201-292-9
; Sequence 9, Application US/10201292
; Publication No. US20030144193A1
; GENERAL INFORMATION:
; APPLICANT: James B. ROTHMAN
; APPLICANT: Theresa L. O'KEEFE
```

```
; APPLICANT: Engin OZKANAK
; APPLICANT: Judith J. HEALEY
; TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
; FILE REFERENCE: 7853-253-999
; CURRENT APPLICATION NUMBER: US/10/201.292
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1650
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-201-292-9

Query Match      86.1%; Score 950.8; DB 15; Length 1650;
Best Local Similarity 99.3%; Pred. No. 7.3e-300;
Matches 955; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATGGCCACGGCGGAGAGAGCCCTCGGCATCGGCTTCCAGTGGCTCTCTTGGCCACT 60
DB 1 ATGGCCACGGCGGAGAGAGCCCTCGGCATCGGCTTCCAGTGGCTCTCTTGGCCACT 60
QY 61 CTGGTCTCATCTGCGCGGCAAGGGGACGACAGAGAGATGAGGGGCTCCAGCTGTAC 120
DB 61 CTGGTCTCATCTGCGCGGCAAGGGGACGACAGAGAGATGAGGGGCTCCAGCTGTAC 120
QY 121 GCGGATTTGACCTGTACTTCACTTTTGGACAAATCAGGAAGTGTCTGACCACTGGAAT 180
DB 121 GCGGATTTGACCTGTACTTCACTTTTGGACAAATCAGGAAGTGTCTGACCACTGGAAT 180
QY 181 GAAATCTATTACTTGTGGACAGTTGGCTCAAAATTCATCAGGCCACAGTTGAGAAATG 240
DB 181 GAAATCTATTACTTGTGGACAGTTGGCTCAAAATTCATCAGGCCACAGTTGAGAAATG 240
QY 241 TCCTTTATTTGTTTCTCCACCCGAGAAACACCTTAAATGAAATGAAATGAAATGAAATG 300
DB 241 TCCTTTATTTGTTTCTCCACCCGAGAAACACCTTAAATGAAATGAAATGAAATGAAATG 300
QY 301 CAATTCCTTCAAGGCTTGAAGAACTCCAGAAAGTTCTGCGAGAGAGACCTTACATG 360
DB 301 CAATTCCTTCAAGGCTTGAAGAACTCCAGAAAGTTCTGCGAGAGAGACCTTACATG 360
QY 361 CATGAAGATTTGAAGGGCCAGTGAAGATTTATTAATGAAACAGACAAGGGTACAG 420
DB 361 CATGAAGATTTGAAGGGCCAGTGAAGATTTATTAATGAAACAGACAAGGGTACAG 420
QY 421 ACAGCAGCGTCTCATCTGCTTGAATGAGAACTTCATGAAGATCTCTTTTCTAT 480
DB 421 ACAGCAGCGTCTCATCTGCTTGAATGAGAACTTCATGAAGATCTCTTTTCTAT 480
QY 481 TCAGAGAGAGAGGCTTAATAGGCTCTGAGATCTTGGTCAATTTGTTACTGTGTGTG 540
DB 481 TCAGAGAGAGAGGCTTAATAGGCTCTGAGATCTTGGTCAATTTGTTACTGTGTGTG 540
QY 541 AAAATTTCAATGAGACACAGCTGCGGATTTGGGACAGTAAGATCATGTGTTCC 600
DB 541 AAAATTTCAATGAGACACAGCTGCGGATTTGGGACAGTAAGATCATGTGTTCC 600
QY 601 GTGAATGACGGCTTCAAGCTCTGCAAGGATCATCACTCAATTTTGAAGAGTCTGCG 660
DB 601 GTGAATGACGGCTTCAAGCTCTGCAAGGATCATCACTCAATTTTGAAGAGTCTGCG 660
QY 661 ATCGAAATTTCTAGCAGCTGAAACATCCATATGTGCAAGAGTCACTTCAAGTGTG 720
DB 661 ATCGAAATTTCTAGCAGCTGAAACATCCATATGTGCAAGAGTCACTTCAAGTGTG 720
QY 721 GTGAGAGAAACGGCTTCCGACATGCGCGCAACGTGACAGGGTCTCTGCAAGTCTCAAG 780
DB 721 GTGAGAGAAACGGCTTCCGACATGCGCGCAACGTGACAGGGTCTCTGCAAGTCTCAAG 780
QY 781 ATCAATGACCTGCTGACACTCAATGAGAAAGCTTTCTGTGAGAAACATTAATTAATG 840
DB 781 ATCAATGACCTGCTGACACTCAATGAGAAAGCTTTCTGTGAGAAACATTAATTAATG 840
```

Qy	841	TGTCAGGCGCCTTATCTTAAAGAAGTTGGCATGAAGCTGCACTCCAGTCAGCATGAAC	900
Db	841	TGTCAGGCGCCTTATCTTAAAGAAGTTGGCATGAAGCTGCACTCCAGTCAGCATGAAC	900
Qy	901	GATGGCCTCTCTTTTATCTCCAGTTCTGTCAATCAACCAACACACTGTTCTGACGGT	960
Db	901	GATGGCCTCTCTTTTATCTCCAGTTCTGTCAATCAACCAACACACTGTTCTGACGGT	960
Qy	961	TC 962	
Db	961	TC 962	

Search completed: June 14, 2005, 18:37:32
 Job time : 786 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2005, 13:05:23 ; Search time 4251 Seconds
(without alignment)
9885.428 Million cell updates/sec

Title: US-09-970-076-1_COPY_104_1207

Perfect score: 1104
Sequence: 1 atggccacggcgagcgag.....gtgaggaataataaaaa 1104

Scoring table: IDENTITY NUC
Gapop 10'-0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_hc3: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_g881: *
9: gb_g882: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	949.4	86.0	1957	3	CR601167 full-length
2	853.4	77.3	3161	3	AK031465 Mus muscu
3	689.2	62.0	1006	1	AL542724 AL542724
4	684.6	62.0	878	4	BT823853 603039031
5	681.4	61.7	1614	3	AK013005 Mus muscu
6	662	60.0	1054	4	BM800898 AGENCOURT
7	658.8	59.7	833	7	CO245219 AGENCOURT
8	637.8	57.8	925	2	BE741333 BE741333
9	585	53.0	963	4	BG326444 602425082
10	582.6	52.8	820	4	BG281561 602402412
11	552	50.0	641	7	CF132432 CF132432
12	541.4	49.0	780	7	CF126791 CF126791
13	534	48.4	823	5	BU131302 603117191
14	533.4	48.3	564	6	CD664532 92B4 IMR3
15	531.4	48.1	877	5	CD664532 92B4 IMR3
16	526.2	47.7	682	7	BX364961 BX364961
17	520.2	47.1	727	6	CF540285 CF540285
18	507.4	46.0	600	6	CA324840 CA324840
19	498.4	45.1	522	6	CD664533 91G4 IMR3
20	482	43.7	964	4	CD664531 91H7 IMR3
21	467.4	42.3	3178	3	BG281831 BG281831
22	466.6	42.3	473	7	AK048637 AK048637
23	464.6	42.1	642	7	CY028654 CY028654
24	461	41.8	600	7	CN461036 CN461036

C	25	451.6	40.9	643	5	BU632284	BU632284	UT-H-FE1-
	26	447.4	40.5	638	2	BE369415	BE369415	601220816
	27	446.8	40.5	582	2	BK470942	BK470942	DKF2P6861
	28	441.8	40.0	805	6	CB990959	CB990959	AGENCOURT
	29	437.4	38.7	793	5	BX369960	BX369960	BX369960
	30	421.2	38.2	867	5	BX337000	BX337000	BX337000
	31	420.4	38.1	513	2	BE146084	BE146084	MRO-HT020
	32	420.4	37.7	602	5	BP456257	BP456257	BP456257
	33	416.4	37.1	524	6	CA945449	CA945449	UT-M-F00-
	34	409.6	37.0	775	7	CK787494	CK787494	AGENCOURT
	35	408.2	36.2	465	2	BE145958	BE145958	MRO-HT020
	36	399.8	34.9	487	2	BE146355	BE146355	MRO-HT020
	37	382	34.6	775	5	B0363718	B0363718	603788453
	38	381.2	34.5	539	1	A1466645	A1466645	mg81d04.Y
	39	376.4	34.1	693	4	BG323538	BG323538	602422195
	40	373.6	33.8	554	1	A1084806	A1084806	0280d12.X
	41	373.6	32.1	546	6	CB519323	CB519323	UT-M-GH0-
	42	354.4	30.7	592	7	CF728459	CF728459	UT-M-HB0-
	43	348.6	30.3	736	6	CB724689	CB724689	EST0528 R
	44	334.2	29.9	356	2	BE146135	BE146135	MRO-HT020
	45	329.8						

ALIGNMENTS

RESULT 1	CR601167	1957 bp	MRNA	linear	HTC 21-JUN-2004
LOCUS	CR601167				
DEFINITION	full-length cDNA clone CSOD1035YA13 of Placenta Cot 25-normalized				
ACCESSION	CR601167				
VERSION	CR601167.1	GI:50481974			
KEYWORDS	HTC; CNS/LT cDNA				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
AUTHORS	Li, W.-B., Gruber, C., Jessee, J., and Polayes, D.				
TITLE	Full-length cDNA libraries and normalized				
JOURNAL	Unpublished				
REMARK	Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue				
REFERENCE	2 (bases 1 to 1957)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage ; BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr				
COMMENT	- Web : www.genoscope.cns.fr) 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.				
FEATURES	Location/Qualifiers				
SOURCE	1..1957				
ORIGIN	/organism="Homo sapiens"				
	/mol_type="mrna"				
	/db_xref="taxon:9606"				
	/clone="CSOD1035YA13"				
	/issue_type="Placenta Cot 25-normalized"				
	/plasmid="pCMVSPORT_6"				
Query Match	86.0% ; Score 949.4 ; DB 3 ; Length 1957 ;				
Best Local Similarity	99.9% ; Pred. No. 2.1e-270 ;				
Matches 950 ; Conservative 0 ; Mismatches 1 ; Indels 0 ; Gaps 0 ;					
Qy	1 ATGGCCACGGCGAGAGCCCTTGGCATGGCTTCCAGTGGCTCTCTTTGGCCACT 60				
Db	195 ATGGCCACGGCGAGAGCCCTTGGCATGGCTTCCAGTGGCTCTCTTTGGCCACT 254				

QY 61 CTGGTGTCTCATCTGCGCGGCAAGGGGAGCGAGGGAGATGGGGGCTCCAGCTGTCTAC 120
DB 255 CTGGTGTCTCATCTGCGCGGCAAGGGGAGCGAGGGAGATGGGGGCTCCAGCTGTCTAC 314
QY 121 GGGCGATTGACCTGTACTTATTTTGGACAATTCAGAAAGTGTGTGCACCACTGGAAT 180
DB 315 GGGCGATTGACCTGTACTTATTTTGGACAATTCAGAAAGTGTGTGCACCACTGGAAT 374
QY 181 GAATTCATTTACTTGTGGAGACAGTTGGCTCAAAATTCATCAGCCCAAGTTAGAAATG 240
DB 375 GAATTCATTTACTTGTGGAGACAGTTGGCTCAAAATTCATCAGCCCAAGTTAGAAATG 434
QY 241 TCCTTTATTTGTTTTCTCCACCCGAGAGAACACTTAAATGAAATGACAGAGAGAGAA 300
DB 435 TCCTTTATTTGTTTTCTCCACCCGAGAGAACACTTAAATGAAATGACAGAGAGAGAA 494
QY 301 CAAATTCGCTCAAGGCTTGAAGAACTCCAGAAAGTTCTCCAGAGAGAGACACTTACATG 360
DB 495 CAAATTCGCTCAAGGCTTGAAGAACTCCAGAAAGTTCTCCAGAGAGAGACACTTACATG 554
QY 361 CATGAAGATTGAAAGGGCCAGAGAGAGATTTATTAAGAAACAGACAAAGGGTACAGG 420
DB 555 CATGAAGATTGAAAGGGCCAGAGAGAGATTTATTAAGAAACAGACAAAGGGTACAGG 614
QY 421 ACAGCCAGCGTCATCATTTGCTTTGACTGATGAGAGAACTCCATGAAGATCTCTTTTCTAT 480
DB 615 ACAGCCAGCGTCATCATTTGCTTTGACTGATGAGAGAACTCCATGAAGATCTCTTTTCTAT 674
QY 481 TCAGAGAGAGAGGCTTAATAGCTCTCGAGATCTTGGTGCATTTGTTACTGTGTGGTGTG 540
DB 675 TCAGAGAGAGAGGCTTAATAGCTCTCGAGATCTTGGTGCATTTGTTACTGTGTGGTGTG 734
QY 541 AAAGATTTCAATGAGACAGAGCTGGCCGGATGGCGAGATGGCGAGATTAAGATCATGTGTTCCC 600
DB 735 AAAGATTTCAATGAGACAGAGCTGGCCGGATGGCGAGATTAAGATCATGTGTTCCC 794
QY 601 GTGATGACGGCTTTCAAGGCTCTCGAAGAGCATATCCACTCAATTTTGAAGAGTCTCTGC 660
DB 795 GTGATGACGGCTTTCAAGGCTCTCGAAGAGCATATCCACTCAATTTTGAAGAGTCTCTGC 854
QY 661 ATCGAATTTCTAGACAGCTGAACCATCCACATATGTGAGAGAGTCAATTTCAAGTTGTC 720
DB 855 ATCGAATTTCTAGACAGCTGAACCATCCACATATGTGAGAGAGTCAATTTCAAGTTGTC 914
QY 721 GTGAGAGAGAAAGGCTTCGACATGCGCGCAACGGTGCAGAGGTCCTTGCAGCTTCAG 780
DB 915 GTGAGAGAGAAAGGCTTCGACATGCGCGCAACGGTGCAGAGGTCCTTGCAGCTTCAG 974
QY 781 ATCAATGACCTGCTCACACTCAATGAGAAAGCCCTTTTCTGTGAGAGACACTTATTTACTG 840
DB 975 ATCAATGACCTGCTCACACTCAATGAGAAAGCCCTTTTCTGTGAGAGACTTATTTACTG 1034
QY 841 TGTCCAGGCGCTATTTTAAAGAAAGTGTGGCATGAAGTGCATCTCCAGTCAAGATGAAC 900
DB 1035 TGTCCAGGCGCTATTTTAAAGAAAGTGTGGCATGAAGTGCATCTCCAGTCAAGATGAAC 1094
QY 901 GATGGCCTCTGCTTTATCTCCAGTTCTGTGATCATCACACACACACTGT 951
DB 1095 GATGGCCTCTGCTTTATCTCCAGTTCTGTGATCATCACACACACACTGT 1145

RESULT 2

AK031465

LOCUS

AK031465 3161 bp mRNA linear HTC 03-APR-2004
Mus musculus 13 days embryo male testis cDNA, RIKEN full-length
enriched library, clone:6030436p19 product:ANTPAX TOXIN RECEPTOR
PRECURSOR (TUMOR ENDOTHELIAL MARKER 8) homolog [Mus musculus], full
insert sequence.

ACCESSION

AK031465 AK031465.1 GI:26082340

VERSION

KEYWORDS

HTC; CAP trapper.
Mus musculus (house mouse)

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99279253

10349636

2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.Normalization and subraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20499374

11042159

3

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Hazada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsuda, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, U.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

20530913

11076861

4

The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.

Functional annotation of a full-length mouse cDNA collection

Nature 409, 685-690 (2001)

5

The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

6

(bases 1 to 3161)

Adachi, U., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numata, R., Ono, M., Ohata, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shitaki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
Location/Qualifiers1. 3161
/organism="Mus musculus"

COMMENT

AK031465

LOCUS

AK031465 3161 bp mRNA linear HTC 03-APR-2004
Mus musculus 13 days embryo male testis cDNA, RIKEN full-length
enriched library, clone:6030436p19 product:ANTPAX TOXIN RECEPTOR
PRECURSOR (TUMOR ENDOTHELIAL MARKER 8) homolog [Mus musculus], full
insert sequence.

ACCESSION

AK031465 AK031465.1 GI:26082340

VERSION

KEYWORDS

HTC; CAP trapper.
Mus musculus (house mouse)

```
/mol_type="mRNA"
/strain="CS7BL/6J"
/db_xref="PANTOM,DB:6030436P19"
/db_xref="taxon:10090"
/clone="6030436P19"
/sex="male"
/classuse_type="leastic"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="13 days embryo"
1..3161
/notes="ANTHEX TOXIN RECEPTOR PRECURSOR (TUMOR ENDOTHELIAL
MARKER 8) homolog [mus musculus] (SWISSPROT Q99C52,
evidence: PASTY, 100%ID, 84.6%length, match=1428)"

ORIGIN
Query Match      77.3%; Score 853.4; DB 3; Length 3161;
Best Local Similarity 87.3%; Pred. No. 9.4e-242;
Matches 935; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

24  CCTGGCAGATCGGCTTCCAGTGGCTCTTTGGCCACTGTGCTCATCTGCGCCGGGCA 83
Db    341  CCTGGGTGCGGGGCTGCGGGGACTCTGCGTGGCTGCACTGCTGCTGCTGCGCGGCA 400
Qy    84  AGGGGGACGCGAGAGATGGGGGTCCAGCTGCTACGCGCGGATTGACCTGTAATTAT 143
Db    401  CGGGGGCGCGCGGAGATGGGGGACCACTTGTCTACGAGGATTCGACCTTAATTCTAT 460
Qy    144  TTTGGACAAATCAGAGATGTGCTGACACCACTGGAATCTATTACTTTGTGGAACA 203
Db    461  CTTGGACAAATCAGAGATGTGCTGACACCACTGGAATCTATTACTTTGTGGAACA 520
Qy    204  GTTGGCTCAAAATTCATCAGCCCAAGTGAAGATCTTTATTTGTTTCTCCACCG 263
Db    521  GTTGGCTCATAGATTCATCAGCCCAAGTGAAGATCTTTATTTGTTTCTCTACTCG 580
Qy    264  AGGAACAACCTTAATGAAGCTGACAGAGAGAGAACTCCGTCAAGGCTTGAAGA 323
Db    581  AGGGACAACTTAAATGAAGCTGAGAGAGAGAGAACTCCGACAAAGGCTTGAAGA 640
Qy    324  ACTCCAGAAAGTTTGCAGAGAGAGACCTTAATCATGATGAAGATTGAAAGGGCCAG 383
Db    641  GCTCCAGAAAGTTTGCAGAGAGAGACCTTAATCATGAGAGATTGCAAGGGCCAG 700
Qy    384  TGAGCAGATTATATGAAAAAGAGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAG 443
Db    701  TGAGCAGATTATATGAGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 760
Qy    444  GACTGATGAGAGATCTCAATGAAGATCTTTTCTTATTCAGAGAGAGAGAGAGAG 503
Db    761  GACCGATGGGAGAGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 820
Qy    504  TCGAGATCTGGTGCATATGTTTATCTGTGTTGTGTAAGAGATTCAATGAGACAG 563
Db    821  CCGAGACCTTGGTGCATATGTTTATCTGTGTTGTGTAAGAGATTCAATGAGAG 880
Qy    564  GGCCCGGATTCGCGAGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 623
Db    881  GGCTCGGATTCGCGAGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 940
Qy    624  GCAAGGATCATCATCATCATTTTGAAGAGATCTGATGAGAGAGAGAGAGAGAG 683
Db    941  CCAAGGATCATCATCATCATTTTGAAGAGATCTGATGAGAGAGAGAGAGAGAG 1000
Qy    684  ATCCACCATATGTCAGAGAGAGATTCATGATGTCGTCGAGAGAGAGAGAGAGAG 743
Db    1001  ATCCACCATATGTCGAGAGAGAGATTCATGATGTCGTCGAGAGAGAGAGAGAG 1060
Qy    744  TGCCCGCAACGTGAGAGAGAGTCTCTGAGAGAGTTCAGATGAGATGAGATGAG 803
Db    1061  TGCCCGCAATGAGAGAGAGTCTCTGAGAGTTCAGATGAGATGAGATGAGATGAG 1120
Qy    804  TGAAGAGCCCTTTTCTGTGAGAGAGATTTATTTCTGTGTCAGAGAGAGAGAG 863
```

```
Db    1121  TGAGAGAGCCCTTCTGTGAGAGAGACTTATTTCTGTGCTGCCAGACCAATCTGAAAGA 1180
Qy    864  AGTTGGATGAAGAGCTGACCTCAGTTCAGATGAGAGAGAGAGAGAGAGAGAG 923
Db    1181  AGTTGGATGAAGAGCTGACCTCAGTTCAGATGAGAGAGAGAGAGAGAGAGAG 1240
Qy    924  TTCTGTATCATCATCAGACACACACACTGTTGACAGGTTTCCATCTGGCCATG 983
Db    1241  TTCTGTATCATCATCAGACACACACACTGTTGACAGAGAGAGAGAGAGAGAG 1300
Qy    984  GATCTGTCCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1043
Db    1301  GATCTGTCCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1360
Qy    1044  TGTGATATCAAGAGAGAGTCCCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1094
Db    1361  AGTATATCAAGAGAGAGTCCCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1411

RESULT 3
AL542724
LOCUS
DEFINITION
5-PRIME, mRNA sequence.
ACCESSION
AL542724
VERSION
AL542724.3 GI:45718299
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1006)
AUTHORS
L.M.B., Gruber,C., Jessee,J., and Polayres,D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished (2001)
COMMENT
On Feb 15, 2001 this sequence version replaced gi:30548161.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Creteil, CP 5706 - 91057 Evry cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 1660.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?b=CS0DE011CF050P16c=1660.f.
Location/Qualifiers
1..1006
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE011YL09"
/classuse_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT 6, 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and EcoRV sites of the
pCMVSPORT 6 vector. Library was not normalized."

ORIGIN
Query Match      62.4%; Score 689.2; DB 1; Length 1006;
Best Local Similarity 99.0%; Pred. No. 4.3e-193;
Matches 703; Conservative 1; Mismatches 4; Indels 2; Gaps 1;

1  ATGGCCAGCGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
Db    293  ATGGCCAGCGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 352
Qy    61  CTGGTGTCTATCTGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
```


Db 353 CTGGTCTCATCTGCGCGGCAAGGGGACGACGAGGAGATGGGGGTCCAGCCTGCTAC 412
Qy 121 GGGGATTTGACCGTGTCTTCTATTTTGACAAATCAGAAAGTGTGCACTGCAAT 180
Db 413 GGGGATTTGACCGTGTCTTCTATTTTGACAAATCAGAAAGTGTGCACTGCAAT 472
Qy 181 GAAATCTATTACTTTGGGAAACAGTTGGCTCACAATTTCACTACGCCACAGTTGAGATG 240
Db 473 GAAATCTATTACTTTGGGAAACAGTTGGCTCACAATTTCACTACGCCACAGTTGAGATG 532
Qy 241 TCCTTTATTTGTTTCTCCACCCGAGAAACAATTAAATGAAATGACAGAAACAGAGAA 300
Db 533 TCCTTTATTTGTTTCTCCACCCGAGAAACAATTAAATGAAATGACAGAAACAGAGAA 592
Qy 301 CAAATCCGTCAAGGCTTAAAGAACTCCAGAAATTTCTCCAGGAGAGACACTTACATG 360
Db 593 CAAATCCGTCAAGGCTTAAAGAACTCCAGAAATTTCTCCAGGAGAGACACTTACATG 652
Qy 361 CATGAGGATTTGAAAGGCGCAGTGAAGCAATTTATTAAGAAACAGACAGGTTACAGG 420
Db 653 CATGAGGATTTGAAAGGCGCAGTGAAGCAATTTATTAAGAAACAGACAGGTTACAGG 712
Qy 421 ACAGCCAGCGTCATCATTTGTTTGAATGAGAAATCCATGAAGATCTCTTTTCTAT 480
Db 713 ACAGCCAGCGTCATCATTTGTTTGAATGAGAAATCCATGAAGATCTCTTTTCTAT 772
Qy 481 TCAGAGAGGAGGCTTAATAGGCTCGAATCTGTGTGCAATTTTCTGTTGTTGTTG 540
Db 773 TCAGAGAGGAGGCTTAATAGGCTCGAATCTGTGTGCAATTTTCTGTTGTTGTTG 832
Qy 541 AAAGATTTCAATGAGACACAGCTGCGCCGAGTTGCGGACAGTAAGATCATGTGTTCC 600
Db 833 AAAGATTTCAATGAGACACAGCTGCGCCGAGTTGCGGACAGTAAGATCATGTGTTCC 892
Qy 601 GTGAATGAGCGCTTTCAGGCTCTGCAAGCATCATCCACTCAATTTTGAAGAATCTCTC 660
Db 893 GTGAATGAGCGCTTTCAGGCTCTGCAAGCATCATCCACTCAATTTTGAAGAATCTCTC 952
Qy 661 ATCGAATTTCTAGAGAGTGAACATCCACATATGTGAGAGAGTCACT 710
Db 953 ATCGAATTTCTAGAGAGTGAACATCCACATATGTGAGAGAGTCACT 1000

RESULT 4
BI823853 878 bp mRNA linear EST 04-OCT-2001
LOCUS 603038031F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5179974 5',
DEFINITION mRNA sequence.
ACCESSION BI823853
VERSION BI823853.1 GI:15935416
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 878)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: csapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHAM1448 row: n column: 07
High quality sequence stop: 804.
Location/Qualifiers
1..878
/organism="Homo sapiens"

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5179974"
/lab_host="DH10B"
/clone_lib="NIH_MGC_115"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 62.0%; Score 684.6; DB 4; Length 878;
Best local similarity 95.1%; Pred. No. 9,4e-192;
Matches 783; Conservative 0; Mismatches 29; Indels 11; Gaps 7;
Qy 1 ATGGCCACGGGAGGAGGAGGCGCTCGGATGGCTCCAGTGGCTCTTTGGCCACT 60
Db 57 ATGGCCACGGGAGGAGGAGGCGCTCGGATGGCTCCAGTGGCTCTTTGGCCACT 116
Qy 61 CTGTGCTCATCTGCGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
Db 117 CTGTGCTCATCTGCGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 176
Qy 121 GGGGATTTGACCTGTACTTCAATTTTGAACAATGAGAAATGCTGTGCACTGGAAT 180
Db 177 GGGGATTTGACCTGTACTTCAATTTTGAACAATGAGAAATGCTGTGCACTGGAAT 236
Qy 181 GAAATCTATTACTTTGGGAAACAGTTGGCTCAAAATTCATCAGCCACAGTTGAGATG 240
Db 237 GAAATCTATTACTTTGGGAAACAGTTGGCTCAAAATTCATCAGCCACAGTTGAGATG 296
Qy 241 TCCTTTATTTGTTTCTCCACCCGAGAAACAATTAAATGAAATGACAGAAACAGAGAA 300
Db 297 TCCTTTATTTGTTTCTCCACCCGAGAAACAATTAAATGAAATGACAGAAACAGAGAA 356
Qy 301 CAAATCCGTCAAGGCTTGAAGAACTCCAGAAAGTTCTGCGAGAGAGACACTTACATG 360
Db 357 CAAATCCGTCAAGGCTTGAAGAACTCCAGAAAGTTCTGCGAGAGAGACACTTACATG 416
Qy 361 CATGAGGATTTGAAAGGCGCAGTGAAGCAATTTATGAAACAGACAGGTTACAGG 420
Db 417 CATGAGGATTTGAAAGGCGCAGTGAAGCAATTTATGAAACAGACAGGTTACAGG 476
Qy 421 ACAGCCAGCGTCATCATTTGTTTGAATGAGAAATCCATGAAGATCTCTTTTCTAT 480
Db 477 ACAGCCAGCGTCATCATTTGTTTGAATGAGAAATCCATGAAGATCTCTTTTCTAT 536
Qy 481 TCAGAGAGGAGGCTTAATAGGCTCGAATCTGTGTGCAATTTTCTGTTGTTGTTG 540
Db 537 TCAGAGAGGAGGCTTAATAGGCTCGAATCTGTGTGCAATTTTCTGTTGTTGTTG 595
Qy 541 AAAGATTTCAATGAGACACAGCTGCGCCGAGTTGCGGACAGTAAGATCATGTGTTCC 600
Db 596 AAAGATTTCAATGAGACACAGCTGCGCCGAGTTGCGGACAGTAAGATCATGTGTTCC 655
Qy 601 GTGAATGAGCGCTTTCAGGCTCTGCAAGCATCATCACTCAATTTTGAAGAATCTCTC 659
Db 656 GTGAATGAGCGCTTTCAGGCTCTGCAAGCATCATCACTCAATTTTGAAGAATCTCTC 715
Qy 660 CATC-GAAATTTCTAGAGCTGAACCATCCATATGTGCAAGAGATC-ATTCAAGTT 717
Db 716 CATCGGAAATTTCTAGAGCTGAACCATCCATATGTGCAAGAGATCATTCAAGTT 775
Qy 718 GTGTGAGAG--AAAGCGCTTCGACATGCGG--CAAGTGGACAGG--TCCTCTG 770
Db 776 GTGTGAGAGAAACGGGTTTCCGACATGCGGCAAAAGTGGACAGGCTCTTGG 835

QY 771 CAGCTTCAAGATCATGACTCGGTACACTCATGAGAACCC 813
Db 836 AGGTTCCAGATCAGACTTCGGTACACTTCAATGAAAGGCC 878

RESULT 5
AK013005

LOCUS AK013005 1614 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length
enriched library, clone:2810405N18 product:ANTHRAX TOXIN RECEPTOR
PRECUSOR (TUMOR ENDOTHELIAL MARKER 8) homolog [Mus musculus], full
insert sequence.

ACCESSION AK013005
VERSION AK013005.1 GI:12850099
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99379253
PUBMED 10349636

REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
TITLE Itch, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalizaton and subtraction of cap-trapper-selected cDNAs to
JOURNAL prepare full-length cDNA libraries for rapid discovery of new genes
MEDLINE Genome Res. 10 (10), 1617-1630 (2000)
PUBMED 20499374
REFERENCE 11042159

REFERENCE 3
AUTHORS Shibata, K., Itch, M., Aizawa, K., Nageoka, S., Sasaki, N., Carninci, P.,
TITLE Komuro, H., Akiyama, J., Nishi, K., Kitsuana, T., Iashiro, H., Itch, M.,
Sun, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujitake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanishi, M.,
Okazaki, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
TITLE RIKEN Integrated sequence analysis (RISA) system--384-format
JOURNAL sequencing pipeline with 384 multicapillary sequencer
MEDLINE Genome Res. 10 (11), 1757-1771 (2000)
PUBMED 11076861

REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
TITLE FANTOM Consortium.
JOURNAL Functional annotation of a full-length mouse cDNA collection
MEDLINE Nature 409, 685-690 (2001)
PUBMED 11076861

REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
TITLE Group Phase I & II Team.
JOURNAL Analysis of the mouse transcriptome based on functional annotation
MEDLINE of 60,770 full-length cDNAs
PUBMED Nature 420, 563-573 (2002)

REFERENCE 6
AUTHORS (bases 1 to 1614)
TITLE Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itch, M., Izawa, M.,
Kabukawa, T., Kato, H., Imotani, K., Ishii, Y., Kono, H., Konda, M.,
Koyra, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
Nemura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sogabe, Y.,
Shukui, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome

COMMENT Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
Please visit our web site (http://genome.gsc.riken.jp/) for further
details.

FEATURES
source
location/qualifiers
1..1614
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:2810405N18"
/db_xref="taxon.10090"
/clone="2810405N18"
/tissue type="whole body"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev stage="10, 11 days embryo"
11..1306
/note="unnamed protein product; ANTHRAX TOXIN RECEPTOR
PRECUSOR (TUMOR ENDOTHELIAL MARKER 8) homolog [Mus
musculus] (SWISSPROT|Q9C252, evidence: FASTV, 100%ID,
84.6%length, match=1428)
putative"
/codon_start=1
/protein_id="PAB28501.1"
/db_xref="GI:12850100"
/translation="MKLTEDREBQIRGLLELQKVLPGDGYTMGEFRASRQIYENS
QGRYASVIALTDGELIHDLFFYSERENARSRLGAIYCVGVKDNFENQDLAIAS
KDHYFVNDFQALOGIILHDLILKSKCEIETACGSPFVVVNGDFRARIY
DRVLCSPKIDSVTLNKPRAVEDYLLCPALIKYGMKAAIOLVSNDDGIFSSSV
IITTHCSGDSIATIALIVPFLIALALMHPPLCTVITIKERPPVRSSEEDD
GLPKKMPYVDASTYGGVGGIKRMVKNMGESGTEGAKLEAKARVAVMEQEYS
FPEPRNINNNRRRSPSPKYSPIKGLDALVTLRKGYSRVSVRPPQPDTRRGWR
LTICIGSKHYHPRHDKGPEPTPLKQAMFSSFLERAFQ"
1592..1597
/note="putative"
1614
/note="putative"

ORIGIN
polyA_signal
polyA_site
polyA_site
Query Match 61.7%; Score 681.4; DB 3; Length 1614;
Best Local Similarity 89.0%; Pred. No. 1.1e-190;
Matches 736; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 268 ACAACCTTAATGAAGTGAACAGAGACAGAAACAAATCCGTCAAGGCTTGAAGAACTC 327
Db 2 ACACTTTAATGAAGTGAACAGAGACAGAGAAACAGATCCGACAGGCTTGAAGAACTC 61

QY 328 CAGAAAGTTCTGCCAGAGAGACACTTACATGATGAAGATTGAAAGGCCAGAGAG 387
Db 62 CAGAAAGTTCTGCCAGAGAGACACTTACATGACAGAGATTCGAGAGGCCAGAGAG 121

QY 388 CAGATTATTATGAAGAACAGAGAGGATCAGAGACAGCGCATCATTTGCTTTGACT 447
Db 122 CAGATTACTRTAGAACAGTCAAGATACAGAGCGGACGTCATCATTCGCTTACG 181

QY 448 GATGAGAACTCCATGAAGATCTTTTCTATTCAAGAGAGGAGGCTAAATAGCTTCGA 507

```
Db 182 GATGGGAGCTGCACAGAGACCTCTTCTTCACTCAGAGAGGAGGCTAACCGATCCCGA 241
Qy 508 GATCTGGTGCATTTGTTTACTGTGTGTGTGTAAGATTTCAATGAGACACAGCTGCGC 567
Db 242 GACCTTGGTGGATTGTTTACTGCGTTGGCGTGAAGATTTCAATGAAACTCAATGGCT 301
Qy 568 CGAATGGCGACAGTAAGATCATGTGTTCCCGTGAATGACGCGTTTCAAGCTCTGCA 627
Db 302 CGAATGGCGACAGTAAGACCAAGCTGTTCTGTGAACGACGCGTTTCAAGCTCTCAA 361
Qy 628 GGCATCATCCCACTCAATTTTGAAGAGTCGATTCGAATTTCTAGACGCTGAACCATCC 687
Db 362 GGCATCATCCCACTCAATTTTGAAGAGTCGATTCGAATTTCTAGCGCGCTGAACCATCC 421
Qy 688 ACCATATGTGACAGAGATCATTTTCAAGTTGCTGTGAGAGAAAAGGCTTCCGATGCC 747
Db 422 ACCATATGTGCGGGAGAGATCTTTTCAAGTGGTGGTAAGAGAAATGGCTTCCGATGCC 481
Qy 748 CGCAACGTGACAGGAGCTCTGACGCTTCAAGATCAATGACTGGTCACTCAATGAG 807
Db 482 CGCAATGTGACAGGAGCTCTGACGCTTCAAAATCAATGACTCACTCAATGAG 541
Qy 808 AAGCCCTTTCTGTGAGAGACATTAATTACTGTGTCCAGGCGCTATCTTAAAGAGTT 867
Db 542 AAGCCCTTTCTGTGAGAGACATTAATTGTGCTGTGCCAGACCAATCTTGAAGAGTT 601
Qy 868 GGCATGAAGGTGACCTCCAGCTCAGATGACATGAGAGCTCTCTTTATCTCCAGTTCT 927
Db 602 GGCATGAAGGTGACCTCCAGCTCAGATGAGATGAGAGAGGCGCTGTCTTATCTCCAGTTCT 661
Qy 928 GTCATCATCACACACACACTGTTCTGACGCTTCACTCTGGCCATGCGCTGTGATC 987
Db 662 GTCATCATCACACACACACTGTTCTGACGCTTCACTCTGGCCATGCTGTGATC 721
Qy 988 CTGTTCCGTCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1047
Db 722 CTGTTCCGTCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 781
Qy 1048 ATTATCAAGAGAGTCCCTCCACCCCTGCGAGAGAGTGAAGAAA 1094
Db 782 ATTATCAAGAGAGTCCCTCCACCCCTGCTGTGAGAGAGTGAAGAAA 828

RESULT 6          1054 bp      mRNA      linear      EST 05-MAR-2002
BM800898          AGENCOURT_6420797 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5541228
LOCUS             5', mRNA sequence.
DEFINITION        BM800898
ACCESSION         BM800898.1 GI:19117721
VERSION           EST.
KEYWORDS          Homo sapiens (human)
SOURCE            Homo sapiens
ORGANISM          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE          1 (baee 1 to 1054)
AUTHORS           NIH-MGC http://mgi.nci.nih.gov/.
TITLE             National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL           Unpublished (1999)
COMMENT           Contact: Robert Strausberg, Ph.D.
                  Email: gsabbs-remail.nih.gov
                  Tissue Procurement: ATCC/DC/DTP
                  cDNA Library Preparation: Life Technologies, Inc.
                  cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
                  DNA Sequencing by: Agencourt Bioscience Corporation
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LIML at:
                  http://image.llnl.gov
                  Plate: L14M1238 row: j column: 13
                  High quality sequence stop: 595.
                  Location/Qualifiers
                    1..1054
                    /organism="Homo sapiens"
```

```
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5541228"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NIH MGC 72"
/note="Organ: skin; Vector: pCMV-Sport6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."
```

ORIGIN

Query Match 60.0%; Score 662; DB 4; Length 1054;
Best Local Similarity 97.5%; Pred. No. 5,4e-185;
Matches 704; Conservative 0; Mismatches 15; Indels 3; Gaps 3;

```
Qy 1 ATGGCCACGGCGGAGCGGAGAGCCCTCGGATCGGCTTCCAGTGGCTCTCTTTGGCCACT 60
Db 97 ATGGCCACGGCGGAGCGGAGAGCCCTCGGATCGGCTTCCAGTGGCTCTCTTTGGCCACT 156
Qy 61 CTGGTGTCTATCTGCGCGGGGCAAGGGGACGAGAGGAGATGGGGTCCAGCTGCTAC 120
Db 157 CTGGTGTCTATCTGCGCGGGGCAAGGGGACGAGAGGAGATGGGGTCCAGCTGCTAC 216
Qy 121 GCGGGAATTTGACCTGTAATCTTATTTTGAAGAAATCAGAAATGTGTGCACTGGAAT 180
Db 217 GCGGGAATTTGACCTGTAATCTTATTTTGAAGAAATCAGAAATGTGTGCACTGGAAT 276
Qy 181 GAAATCATTAATCTTTGGAACAGTTGGCTCACAAATTCATCAGCCACAGTTGAAGATG 240
Db 277 GAAATCATTAATCTTTGGAACAGTTGGCTCACAAATTCATCAGCCACAGTTGAAGATG 336
Qy 241 TCCTTTATTTGTTTCTCCACCGAGAAACCTTAATGAATGACAGAAACAGAA 300
Db 337 TCCTTTATTTGTTTCTCCACCGAGAAACCTTAATGAATGACAGAAACAGAA 396
Qy 301 CAATCCGTCAAGGCTTGAAGAAATCCAGAAATTTCTGCAGAGAGACATTAATG 360
Db 397 CAATCCGTCAAGGCTTGAAGAAATCCAGAAATTTCTGCAGAGAGACATTAATG 456
Qy 361 CATGAAGATTTGAAGAGGCGCAGTGACAGATTTAATGAAAAACAGCAAGGTAACG 420
Db 457 CATGAAGATTTGAAGAGGCGCAGTGACAGATTTAATGAAAAACAGCAAGGTAACG 516
Qy 421 ACAGCCAGCGTCATCATTTGCTTGTGATGATGAGAACTCCATGAAGATCTTTCTAT 480
Db 517 ACAGCCAGCGTCATCATTTGCTTGTGATGATGAGAACTCCATGAAGATCTTTCTAT 576
Qy 481 TCAGAGAGGAGGCTAATAGTCTCCAGATCTTGTCGAATTTGTTACTGTGTGTG 540
Db 577 TCAGAGAGGAGGCTAATAGTCTCCAGATCTTGTCGAATTTGTTACTGTGTGTG 636
Qy 541 AAGATTTCAATGAGACAGAGCTGGCCGAGATTGCGGACAGTAAGATGTTTCCC 600
Db 637 AAGATTTCAATGAGACAGAGCTGGCCGAGATTGCGGACAGTAAGATGTTTCCC 696
Qy 601 GTGAATGACGCTTTCAGGCTCTGCAAGGATCATCCATCAATTTTGAAGAGCTCTGC 660
Db 697 GTGAATGACGCTTTCAGGCTCTGC-AGGATCATCCATCCAGTTTGAAGAGCTCTGC 755
Qy 661 ATCAAAATTTCAAGAGCTGAACCATTCACCAT-ATGCGA-GGAGATCAATTTCAAGT 718
Db 756 ATCAAAATTTCAAGGAGGTGAACATCCCATGATGTGCAAGGAGACATTTCCAGGT 815
Qy 719 TC 720
Db 816 TC 817

RESULT 7          833 bp      mRNA      linear      EST 23-JUN-2004
CO245219          AGENCOURT_26524177 NIH_MGC_212 Homo sapiens cDNA clone
LOCUS             833 bp
DEFINITION        AGENCOURT_26524177 NIH_MGC_212 Homo sapiens cDNA clone
```

ACCESSION IMAGE:30924322 5', mRNA sequence.
 VERSION C0245219
 KEYWORDS C0245219.1 GI:49108057
 SOURCE EST.
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 833)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cga@dbp-remail.nih.gov
 Tissue Procurement: Mary Hendrix
 CDNA Library Preparation: M. Bento Soares, University of Iowa
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Plate: NDAm116 row: 1 column: 11
 High quality sequence start: 6
 High quality sequence stop: 688.
 Location/Qualifiers

FEATURES

1..833
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30924322"
 /tissue_type="Chondrosarcoma Lung Metastasis cell lines"
 /lab_host="DH10B (TI phage resistant)"
 /clone_1lb="NIH_MGC_212"
 /note="Organ: Lung; Vector: pYX-ABC; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I and then cloned directionally into pYX-ABC vector. The library tag sequence located between the Not I site and the polyA tail is GATTAAGCCA. Tissue was provided by Mary Hendrix."

ORIGIN

Query Match 59.7%; Score 658.8; DB 7; Length 833;
 Best Local Similarity 97.9%; Pred. No. 4,4e-184;
 Matches 689; Conservative 0; Mismatches 12; Indels 3; Gaps 2;
 Oy 1 ATGGCCACGGGGAGCGGAGACCCCTCGGCACTCGGCTTCCAGTGGCTCTCTTTGGCCACT 60
 Db 100 ATGGCCACGGGGAGCGGAGACCCCTCGGCACTCGGCTTCCAGTGGCTCTCTTTGGCCACT 159
 Oy 61 CTGGTGCTCATCTGCGCGCGGCAAGGGGAGCGAGGAGATGGGGGTCACGCTGTAC 120
 Db 160 CTGGTGCTCATCTGCGCGCGGCAAGGGGAGCGAGGAGATGGGGGTCACGCTGTAC 219
 Oy 121 GGGCGATTTGACCTGTACTTCTTGGCAAAATCAGAGATGTGTGACCACTGGAAT 180
 Db 220 GGGCGATTTGACCTGTACTTCTTGGCAAAATCAGAGATGTGTGACCACTGGAAT 279
 Oy 181 GAATCTATTACTTTGGGAACAGTTGGCTCAAAATTCATCAGCCCAAGTTGAATG 240
 Db 280 GAATCTATTACTTTGGGAACAGTTGGCTCAAAATTCATCAGCCCAAGTTGAATG 339
 Oy 241 TCCCTTATTGTTTCTCACCGGAGAACACTTAATGAACGACGAGAGACAGAGA 300
 Db 340 TCCCTTATTGTTTCTCACCGGAGAACACTTAATGAACGACGAGAGACAGAGA 399

Oy 301 CAATTCGTCAGAGCCCTAGAAAGAACTCCAGAAAGTTTCCGAGAGAGACACTTACATG 360
 Db 400 CAATTCGTCAGAGCCCTAGAAAGAACTCCAGAAAGTTTCCGAGAGAGACACTTACATG 459
 Oy 361 CATGAAGATTGGAAGGGCCAGTGAAGATTATATGAAGAAACAGACAAAGGTTACAG 420
 Db 460 CATGAAGATTGGAAGGGCCAGTGAAGATTATATGAAGAAACAGACAAAGGTTACAG 519
 Oy 421 ACAGCCAGCGTCATCTGCTTTGACTGATGAGAGAACTCCATGAAGATCTTTTTCAT 480
 Db 520 ACAGCCAGCGTCATCTGCTTTGACTGATGAGAGAACTCCATGAAGATCTTTTTCAT 579
 Oy 481 TCAGAGAGGAGGCTAATAGGCTCGAGATCTTGTCGAATTTGTAAGTGTGTTG 540
 Db 580 TCAGAGAGGAGGCTAATAGGCTCGAGATCTTGTCGAATTTGTAAGTGTGTTG 639
 Oy 541 AAATTTCAATGAGACACAGCTGGCCCGGATTTGCCGACGTAAAGATCATGTGTTCC 600
 Db 640 AAATTTCAATGAGACACAGCTGGCCCGGATTTGCCGACGTAAAGATCATGTGTTCC 699
 Oy 601 GTGAATGACGCTTTGAGGCTGTCAGAGCATCATCTCAATTTGAAGAAGTCTGC 660
 Db 700 GTGAATGACGCTTTGAGGCTGTCAGAGCATCATCTCAATTTGAAGAAGTCTGC 759
 Oy 661 ATCGAAATTTAGAGCTGAACCATCCACCATATGTGCAGAGAGA 704
 Db 760 ATCGAAATTTAGAGCTGAAC--ATCCACATATGTGCAGAGAGA 800

RESULT 8
 BE741333
 LOCUS 725 bp mRNA linear EST 15-SEP-2000
 DEFINITION 60159417P1 NIH_MGC_9 Homo sapiens CDNA clone IMAGE:3948114 5', mRNA sequence.
 ACCESSION BE741333
 VERSION BE741333.1 GI:10155325
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 725)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@dbp-remail.nih.gov
 Tissue Procurement: DCTD/DTP
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: image.llnl.gov
 Plate: LNC610 row: n column: 19
 High quality sequence stop: 711.
 Location/Qualifiers

FEATURES

SOURCE

1..725
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3948114"
 /tissue_type="adenocarcinoma cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_1lb="NIH_MGC_9"
 /note="Organ: ovary; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match: 57.8%; Score 637.8; DB 2; Length 725;
Best Local Similarity 98.8%; Pred. No. 7,3e-178;
Matches 653; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

```
QY 292 GACGAGAAACAATTCCTCAAGGCTTGAAGAACTCCAGAAAGTTCTCCAGAGAGAC 351
DB 1 GACGAGAAACAATTCCTCAAGGCTTGAAGAACTCCAGAAAGTTCTCCAGAGAGAC 60
QY 352 ACTTACATGCAATGAAGATTGAAAGGCGCAGTGAAGAACTTTATTAAGAAACAGCAA 411
DB 61 ACTTACATGCAATGAAGATTGAAAGGCGCAGTGAAGAACTTTATTAAGAAACAGCAA 120
QY 412 GGGTACAGACAGCAGCGTCATCTGCTTTGACTGAGAGAACTCAGTGAAGATCTC 471
DB 121 GGGTACAGACAGCAGCGTCATCTGCTTTGACTGAGAGAACTCAGTGAAGATCTC 180
QY 472 TTTTCTATTCAAGAGGAGGCTTAATAGTCTCGAATCTGGTGCATTTTACTGT 531
DB 181 TTTTCTATTCAAGAGGAGGCTTAATAGTCTCGAATCTGGTGCATTTTACTGT 240
QY 532 GTTGTGTGAAGATTTCAATGAGACAGCTGCGCGGATTTGGGACAGTAAGATCAT 591
DB 241 GTTGTGTGAAGATTTCAATGAGACAGCTGCGCGGATTTGGGACAGTAAGATCAT 300
QY 592 GTGTTCCCGTGAATGACGCGCTTCAAGGCTTGCAGGCACTCATCAATTTTGAAG 651
DB 301 GTGTTCCCGTGAATGACGCGCTTCAAGGCTTGCAGGCACTCATCAATTTTGAAG 360
QY 652 AAGTCTGCATCGAAATTTTCAGCTGAGACCATTCACATATGTGCAAGATCTTT 711
DB 361 AAGTCTGCATCGAAATTTTCAGCTGAGACCATTCACATATGTGCAAGATCTTT 420
QY 712 CAATGTGCTGAGAGGAAAGGCTTCCGACATGCGCGCAAGTGAAGAGGCTCTGCG 771
DB 421 CAATGTGCTGAGAGGAAAGGCTTCCGACATGCGCGCAAGTGAAGAGGCTCTGCG 480
QY 772 AGCTTCAAGATCAATGACTCGGTGACACTCAATGAGAGGCTTTCTGTGGAAGACT 831
DB 481 AGCTTCAAGATCAATGACTCGGTGACACTCAATGAGAGGCTTTCTGTGGAAGACT 540
QY 832 TATTACTGTGTCCAGCGCTTATCTTAA-AGAAAGTTGGCATGAAGTGTGACAGT 890
DB 541 TATTACTGTGTCCAGCGCTTATCTTAA-AGAAAGTTGGCATGAAGTGTGACAGT 600
QY 891 CAGATGAAGATGGCGCTCTTTTATCTCCAGATTTGTGTCATCAACCAACACACTG 950
DB 601 CAGATGAAGATGGCGCTCTTTTATCTCCAGATTTGTGTCATCAACCAACACACTG 660
QY 951 T 951
DB 661 T 661
```

RESULT 9
BG326444 963 bp mRNA linear EST 27-FEB-2001
LOCUS 602425082P1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4563020 5',
DEFINITION mRNA sequence.
ACCESSION BG326444
VERSION BG326444.1 GI:13132881
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 963)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@db-remail.nih.gov
Tissue Procurement: DCTD/DTF

CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCML276 row: c column: 21
High quality sequence stop: 718.
Location/Qualifiers

FEATURES

source
1..963
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4563020"
/tissue_type="renal cell adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH MGC 14"
/note="Organ: kidney; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAAGG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match: 53.0%; Score 585; DB 4; Length 963;
Best Local Similarity 96.3%; Pred. No. 4.2e-162;
Matches 621; Conservative 0; Mismatches 20; Indels 4; Gaps 2;

```
QY 1 ATGGCCACGGCGGAGCGGAGAGGCTTCGCGATCGGCTTCCAGTGCTCTTTGGCCACT 60
DB 114 ATGGCCACGGCGGAGCGGAGAGGCTTCGCGATCGGCTTCCAGTGCTCTTTGGCCACT 173
QY 61 CTGTGTCTATCTGCGCCGGGCAAGGGAGCGAGAGGATGGGCTCAGCTGTAC 120
DB 174 CTGTGTCTATCTGCGCCGGGCAAGGGAGCGAGAGGATGGGCTCAGCTGTAC 233
QY 121 GGGGAAATTTGACCTGTACTTCAATTTTGAACAATCAGAAATGTGTGACCACTGGAAT 180
DB 234 GGGGAAATTTGACCTGTACTTCAATTTTGAACAATCAGAAATGTGTGACCACTGGAAT 293
QY 181 GAAATCTATTACTTGTGGAACAGTTGGCTCACAATTCATCAGCCCACTTGAAGATG 240
DB 294 GAAATCTATTACTTGTGGAACAGTTGGCTCACAATTCATCAGCCCACTTGAAGATG 353
QY 241 TCCCTTATTTGTTTCTCCACCGAGGAACAACCTTAATGAACTGACAGAAACAGAA 300
DB 354 TCCCTTATTTGTTTCTCCACCGAGGAACAACCTTAATGAACTGACAGAAACAGAA 413
QY 301 CAATCCGTCAAGGCTTGAAGAACTCCAGAAAGTTCTGCGAGAGAGACATTACATG 360
DB 414 CAATCCGTCAAGGCTTGAAGAACTCCAGAAAGTTCTGCGAGAGAGACATTACATG 473
QY 414 CAATCCGTCAAGGCTTGAAGAACTCCAGAAAGTTCTGCGAGAGAGACATTACATG 473
QY 361 CATGAAGATTTGAAGGCGCAGTGAAGCAATTTATTAAGAAACAGACAAGGTTACAG 420
DB 474 CATGAAGATTTGAAGGCGCAGTGAAGCAATTTATTAAGAAACAGACAAGGTTACAG 533
QY 421 ACAGCCAGCTCATCANTTGTCTTGAATGAGAACTCCATGAAGATCTTTTCTAT 480
DB 534 ACAGCCAGCTCATCANTTGTCTTGAATGAGAACTCCATGAAGATCTTTTCTAT 593
QY 481 TCAGAGAGAGGCTAATAGGTCTCGAGATCTTGTCATTTGTTACTGTGTGTG 540
DB 594 TCAGAGAGAGGCTAATAGGTCTCGAGATCTTGTCATTTGTTACTGTGTGTG 653
QY 541 AAAGATTTCAATGAGACAGCTGGCCGGATTTGGGACAGTAAGATCATGTGTTC 600
DB 654 AAAGATTTCAATGAGACAGCTGGCC--GGAATGGGACAGTAAGATCATGTGTATCCC 711
QY 601 GTGAATGAGGCTTTCAGGCTCTGCAAGGATCATTCACATCAATT 645
```

DB	712GTGAAATGACGG--TTTAAATCTGGAGAGCATCATTCATTCTTT 754
RESULT 10	
LOCUS	BG281561 820 bp mRNA linear EST 21-FEB-2001
DEFINITION	602402412P1 NIH_MGC_20 Homo sapiens CDNA clone IMAGE:4544691 5',
ACCESSION	BG281561
VERSION	BG281561.1 GI:13030486
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 820)
TITLE	NIH-MGC http://mgs.nci.nih.gov/.
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
COMMENT	Contact: Robert Strauszberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: ATCC/DC/DTF CDNA Library Preparation: Ling Hong/Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov Plate: L10CM128 row: h column: 04 High quality sequence stop: 815. Location/Qualifiers
FEATURES	
SOURCE	1..820 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:4544691" /tissue_type="melanotic melanoma" /lab_host="DH10B (phage-resistant)" /clone_11b="NIH_MGC_20" /notes="Organ: skin; Vector: pOTB7; Site_1: XhoI, site_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-sequences >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
ORIGIN	
Query Match	52.8%; Score 582.6; DB 4; Length 820;
Best Local Similarity	96.1%; Pred. No. 2e-161;
Matches 683; Conservative	0; Mismatches 19; Indels 9; Gaps 8;
1	ATGGCGAAGGCGGAGCGAGAGCCCTCGGCAATCGGCTTCAATGGCTCTCTTTGGCACT 60
112	ATGGCGAAGGCGGAGCGAGAGCCCTCGGCAATCGGCTTCAATGGCTCTCTTTGGCACT 171
61	CTGGTGTCTCATCTGCGCGCGGCAAGGCGGAGCGAGAGATGGGGGTCAAGCTGTCTAC 120
172	CTGGTGTCTCATCTGCGCGCGGCAAGGCGGAGCGAGAGATGGGGGTCAAGCTGTCTAC 231
121	GGCGGATTTGACCTGTACTTCTATTTGGACAATATAGGAAGTGTGTCTGACCATGTGAAT 180
232	GGCGGATTTGACCTGTACTTCTATTTGGACAATATAGGAAGTGTGTCTGACCATGTGAAT 291
181	GAATCTATTAATCTTGTGTGAAGAGTT--GGCTCACAAAATTCAATCAGCCACAGTTGAGAT 239
292	GAATCTATTAATCTTGTGTGAAGAGTTGAGTTGACCAATTTCAATCAGCCACAGTTGAGAT 351
240	GTCTCT--TTAATGTTTTTCTCCAGCCGAGAAACAACCTTAATGAATGACAGAGAAGA 297
352	GTCTCTTAATATATGTTTTCTCCAGCCGAGAAACAACCTTAATGAATGACAGAGAAGA 411
298	GACAAATTCGCTCAAGGCTGAGAAAGTCCAGAAAGTTCTGCCAGAGAGAGACACTTAC 357

Db	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	PUBMED	COMMENT
Db	CP132432	CP132432 641 bp mRNA linear EST 05-AUG-2003												
Db	CP132432	UI-HF-F00-awo-m-16-0-UI.r1 NIH MGC_215 Homo sapiens cDNA clone IMAGE:30560079 5', mRNA sequence.												
Db	CP132432	CP132432												
Db	CP132432.1	GI:33215693												
Db	CP132432	641 bp mRNA linear EST 05-AUG-2003												
Db	CP132432	UI-HF-F00-awo-m-16-0-UI.r1 NIH MGC_215 Homo sapiens cDNA clone IMAGE:30560079 5', mRNA sequence.												
Db	CP132432	CP132432												
Db	CP132432.1	GI:33215693												
Db	CP132432	641 bp mRNA linear EST 05-AUG-2003												
Db	CP132432	UI-HF-F00-awo-m-16-0-UI.r1 NIH MGC_215 Homo sapiens cDNA clone IMAGE:30560079 5', mRNA sequence.												
Db	CP132432	CP132432												
Db	CP132432.1	GI:33215693												
Db	CP132432	641 bp mRNA linear EST 05-AUG-2003												
Db	CP132432	UI-HF-F00-awo-m-16-0-UI.r1 NIH MGC_215 Homo sapiens cDNA clone IMAGE:30560079 5', mRNA sequence.												
Db	CP132432	CP132432												
Db	CP132432.1	GI:33215693												
Db	CP132432	641 bp mRNA linear EST 05-AUG-2003												
Db	CP132432	UI-HF-F00-awo-m-16-0-UI.r1 NIH MGC_215 Homo sapiens cDNA clone IMAGE:30560079 5', mRNA sequence.												
Db	CP132432	CP132432												
Db	CP132432.1	GI:33215693												
Db	CP132432	641 bp mRNA linear EST 05-AUG-2003												
Db	CP132432	UI-HF-F00-awo-m-16-0-UI.r1 NIH MGC_215 Homo sapiens cDNA clone IMAGE:30560079 5', mRNA sequence.												
Db	CP132432	CP132432												
Db	CP132432.1	GI:33215693												
Db	CP132432	641 bp mRNA linear EST 05-AUG-2003												
Db	CP132432	UI-HF-F00-awo-m-16-0-UI.r1 NIH MGC_215 Homo sapiens cDNA clone IMAGE:30560079 5', mRNA sequence.												
Db	CP132432	CP132432												
Db	CP132432.1	GI:33215693												
Db	CP132432	641 bp mRNA linear EST 05-AUG-2003												
Db	CP132432	UI-HF-F00-awo-m-16-0-UI.r1 NIH MGC_215 Homo sapiens cDNA clone IMAGE:30560079 5', mRNA sequence.												
Db	CP132432	CP132432												
Db	CP132432.1	GI:33215693												
Db	CP132432	641 bp mRNA linear EST 05-AUG-2003												
Db	CP132432	UI-HF-F00-awo-m-16-0-UI.r1 NIH MGC_215 Homo sapiens cDNA clone IMAGE:30560079 5', mRNA sequence.												
Db	CP132432	CP132432												
Db	CP132432.1	GI:33215693												
Db	CP132432	641 bp mRNA linear EST 05-AUG-2003												
Db	CP132432	UI-HF-F00-awo-m-16-0-UI.r1 NIH MGC_215 Homo sapiens cDNA clone IMAGE:30560079 5', mRNA sequence.												
Db	CP132432	CP132432												
Db	CP132432.1	GI:33215693												
Db	CP132432	641 bp mRNA linear EST 05-AUG-2003												

1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I and then cloned directionally into pX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is GATTAAGGCCA. Tissue was provided by Mary Hendrix."

ORIGIN

Query Match 50.0%; Score 552; DB 7; Length 641;
Best Local Similarity 99.8%; Pred. No. 2,4e-152;
Matches 563; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 ATGGCCACGGCGGAGGAGAGCCCTGGCATCGGCTTCCAGTGGCTCTCTTTGGCCACT 60
DB 79 ATGGCCACGGCGGAGGAGAGCCCTGGCATCGGCTTCCAGTGGCTCTCTTTGGCCACT 138
QY 61 CTGGTGTCTCATCTGCGCGGCGCAAGGGGAGCGCAGGGAGAGTGGGGTCCAGCTGTCTAC 120
DB 139 CTGGTGTCTCATCTGCGCGGCGCAAGGGGAGCGCAGGGAGAGTGGGGTCCAGCTGTCTAC 198
QY 121 GGGCGATTTGACCTGTACTTCTATTTTGACAAATCAGAAAGTGTGCTGCACTGGAAT 180
DB 199 GGGCGATTTGACCTGTACTTCTATTTTGACAAATCAGAAAGTGTGCTGCACTGGAAT 258
QY 181 GAAATCTATTACTTGTGTGGAAGAGTGGCTCAAAATTCATCAGCCCAAGTTGAGATG 240
DB 259 GAAATCTATTACTTGTGTGGAAGAGTGGCTCAAAATTCATCAGCCCAAGTTGAGATG 318
QY 241 TCGTTTATTTGTTTCTCCACCCGAGAAACAATCTTAATGAAATCTGACAGAAACAGAGAA 300
DB 319 TCGTTTATTTGTTTCTCCACCCGAGAAACAATCTTAATGAAATCTGACAGAAACAGAGAA 378
QY 301 CAAATCGTCAAGGCTTGAAGAACTCCAGAAAGTGTGCGAGAGAGACACTTAATG 360
DB 379 CAAATCGTCAAGGCTTGAAGAACTCCAGAAAGTGTGCGAGAGAGACACTTAATG 438
QY 361 CATGAAGATTTGAAAGGGCCAGTGAAGATTTATATGAAAAACAGACAGGTTACAGG 420
DB 439 CATGAAGATTTGAAAGGGCCAGTGAAGATTTATATGAAAAACAGACAGGTTACAGG 498
QY 421 ACAGCCAGCGTCATTCATTTGATCTGATGAGAACTCCATGAAGATCTCTTTTCTAT 480
DB 499 ACAGCCAGCGTCATTCATTTGATCTGATGAGAACTCCATGAAGATCTCTTTTCTAT 558
QY 481 TCAGAGAGGGAGGCTTAATGAGTCTGAGATCTGTGTCGAATTTTATCTGTTGGTGTG 540
DB 559 TCAGAGAGGGAGGCTTAATGAGTCTGAGATCTGTGTCGAATTTTATCTGTTGGTGTG 618
QY 541 AAAGATTTCAATGAGACACAGCTG 564
DB 619 AAAGATTTCAATGAGACACAGCTG 641

RESULT 12
CF126791 780 bp mRNA linear EST 05-AUG-2003
LOCUS CF126791
DEFINITION UT-HF-BT0-aav-o-15-o-UI.r1 NIH_MGC_214 Homo sapiens cDNA clone
IMAGE:30560510 5', mRNA sequence.

ACCESSION CF126791
VERSION CF126791.1 GI:33204381
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

AUTHORS Bernaldo, M.F., Lennon, G. and Soares, M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene
discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477

PUBMED

8889548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Mary Hendrix
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/humanfl.html>
Seq primer: pX-5.

FEATURES

source
1..780
location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30560510"
/tissue type="Chondrosarcoma Lung Metastasis cell lines"
/lab host="DH10B (T1 phage resistant)"
/clone_1lb="NIH_MGC_214"
/note="Organ: Lung; Vector: pX-Asc; Site_1: EcoR I;
Site_2: Not I; The library was constructed according
Bernaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with Not I and then cloned
directionally into pX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is GATTAAGGCCA. Tissue was provided by Mary Hendrix."

ORIGIN

Query Match 49.0%; Score 541.4; DB 7; Length 780;
Best Local Similarity 99.8%; Pred. No. 3.6e-149;
Matches 542; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCCACGGCGGAGGAGAGCCCTGGCATCGGCTTCCAGTGGCTCTCTTTGGCCACT 60
DB 238 ATGGCCACGGCGGAGGAGAGCCCTGGCATCGGCTTCCAGTGGCTCTCTTTGGCCACT 297
QY 61 CTGGTGTCTCATCTGCGCGGCGCAAGGGGAGCGCAGGGAGATGGGGTCCAGCTGTCTAC 120
DB 298 CTGGTGTCTCATCTGCGCGGCGCAAGGGGAGCGCAGGGAGATGGGGTCCAGCTGTCTAC 357
QY 121 GGGCGATTTGACCTGTACTTCTATTTTGACAAATCAGAAAGTGTGCTGACCACTGGAAT 180
DB 358 GGGCGATTTGACCTGTACTTCTATTTTGACAAATCAGAAAGTGTGCTGACCACTGGAAT 417
QY 181 GAAATCTATTACTTGTGTGGAAGAGTGGCTCAAAATTCATCAGCCCAAGTTGAGATG 240
DB 418 GAAATCTATTACTTGTGTGGAAGAGTGGCTCAAAATTCATCAGCCCAAGTTGAGATG 477
QY 241 TCGTTTATTTGTTTCTCCACCCGAGAAACAATCTTAATGAAATCTGACAGAAACAGAGAA 300
DB 478 TCGTTTATTTGTTTCTCCACCCGAGAAACAATCTTAATGAAATCTGACAGAAACAGAGAA 537
QY 301 CAAATCGTCAAGGCTTGAAGAACTCCAGAAAGTGTGCGAGAGAGACACTTAATG 360
DB 538 CAAATCGTCAAGGCTTGAAGAACTCCAGAAAGTGTGCGAGAGAGACACTTAATG 597
QY 361 CATGAAGATTTGAAAGGGCCAGTGAAGATTTATATGAAAAACAGACAGGTTACAGG 420
DB 598 CATGAAGATTTGAAAGGGCCAGTGAAGATTTATATGAAAAACAGACAGGTTACAGG 657
QY 421 ACAGCCAGCGTCATTCATTTGATCTGATGAGAACTCCATGAAGATCTCTTTTCTAT 480
DB 658 ACAGCCAGCGTCATTCATTTGATCTGATGAGAACTCCATGAAGATCTCTTTTCTAT 717

Thu Jun 16 13:07:15 2005

us-09-970-076-1_copy_104_1207.rst

Page 13

Job time : 4263 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 19:46:02 ; Search time 28.481 Seconds
(without alignments)
773.200 Million cell updates/sec

Title: US-09-970-076-2_COPY_27_321
Perfect score: 1536
Sequence: 1 GCGGRREDGPGACGFDLY.....GLSPFISSVITTTTCSDGS 295

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.*
1: /cgn2_6/ptcdatc/1/1aa/5A.COMB.pep.*
2: /cgn2_6/ptcdatc/1/1aa/5B.COMB.pep.*
3: /cgn2_6/ptcdatc/1/1aa/6A.COMB.pep.*
4: /cgn2_6/ptcdatc/1/1aa/6B.COMB.pep.*
5: /cgn2_6/ptcdatc/1/1aa/PTCUS.COMB.pep.*
6: /cgn2_6/ptcdatc/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	166	10.8	1155	1 US-08-286-889-46	Sequence 46, App1
2	166	10.8	1155	1 US-08-485-618-46	Sequence 46, App1
3	166	10.8	1155	1 US-08-362-652-46	Sequence 46, App1
4	166	10.8	1155	2 US-08-605-672-46	Sequence 46, App1
5	166	10.8	1155	2 US-08-482-293A-46	Sequence 46, App1
6	166	10.8	1155	2 US-08-943-363-46	Sequence 46, App1
7	166	10.8	1155	3 US-09-193-043-46	Sequence 46, App1
8	166	10.8	1155	4 US-09-688-307A-46	Sequence 46, App1
9	166	10.8	1155	4 US-09-350-259-46	Sequence 46, App1
10	166	10.8	1161	1 US-08-485-618-53	Sequence 53, App1
11	166	10.8	1161	2 US-08-605-672-53	Sequence 53, App1
12	166	10.8	1161	2 US-08-482-293A-53	Sequence 53, App1
13	166	10.8	1161	2 US-08-943-363-53	Sequence 53, App1
14	166	10.8	1161	3 US-09-193-043-53	Sequence 53, App1
15	166	10.8	1161	4 US-09-688-307A-53	Sequence 53, App1
16	166	10.8	1161	4 US-09-350-259-53	Sequence 53, App1
17	166	10.8	1161	4 US-08-286-889-37	Sequence 37, App1
18	159.5	10.4	1151	1 US-08-485-618-37	Sequence 37, App1
19	159.5	10.4	1151	1 US-08-605-672-37	Sequence 37, App1
20	159.5	10.4	1151	2 US-08-482-293A-37	Sequence 37, App1
21	159.5	10.4	1151	2 US-08-943-363-37	Sequence 37, App1
22	159.5	10.4	1151	2 US-08-193-043-37	Sequence 37, App1
23	159.5	10.4	1151	3 US-08-688-307A-37	Sequence 37, App1
24	159.5	10.4	1151	4 US-08-350-259-37	Sequence 37, App1
25	159.5	10.4	1151	4 US-08-485-618-55	Sequence 55, App1
26	159.5	10.4	1151	1 US-08-485-618-55	Sequence 55, App1
27	159.5	10.4	1161	1 US-08-485-618-55	Sequence 55, App1

28	159.5	10.4	1161	1 US-08-362-652-55	Sequence 55, App1
29	159.5	10.4	1161	2 US-08-605-672-55	Sequence 55, App1
30	159.5	10.4	1161	2 US-08-482-293A-55	Sequence 55, App1
31	159.5	10.4	1161	2 US-08-943-363-55	Sequence 55, App1
32	159.5	10.4	1161	3 US-09-193-043-55	Sequence 55, App1
33	159.5	10.4	1161	4 US-09-688-307A-55	Sequence 55, App1
34	159.5	10.4	1161	4 US-09-350-259-55	Sequence 55, App1
35	156	10.2	1161	1 US-08-173-497-2	Sequence 2, App1
36	156	10.2	1161	1 US-08-286-889-2	Sequence 2, App1
37	156	10.2	1161	1 US-08-485-618-2	Sequence 2, App1
38	156	10.2	1161	1 US-08-485-618-99	Sequence 99, App1
39	156	10.2	1161	1 US-08-362-652-2	Sequence 2, App1
40	156	10.2	1161	2 US-08-605-672-2	Sequence 2, App1
41	156	10.2	1161	2 US-08-482-293A-2	Sequence 2, App1
42	156	10.2	1161	2 US-08-943-363-2	Sequence 2, App1
43	156	10.2	1161	2 US-08-482-293A-99	Sequence 99, App1
44	156	10.2	1161	2 US-08-943-363-2	Sequence 2, App1
45	156	10.2	1161	2 US-08-943-363-99	Sequence 99, App1

ALIGNMENTS

RESULT 1
US-08-286-889-46
Sequence 46, Application US/08286889
Patent No. 5470953
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Mich
APPLICANT: Van der Vlieten, Monica
TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/286,889
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: P38,659
REFERENCE/DOCKET NUMBER: 27866/32168
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 1155 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULAR TYPE: protein
US-08-286-889-46

Query Match 10.8%; Score 166; DB 1; Length 1155;
Best Local Similarity 28.2%; Pred. No. 6.9e-09;
Matches 66; Conservative 42; Mismatches 96; Indels 30; Gaps 12;
11 PACYG-GFDLYPLDSSGV-LHHNWEIYFVEQLAHKFTSPOLRMSFVFSRGTTLMK 68

Db 144 PEPGQEMDIAFLIDSGSIDQSDFTQMKDFVKALMGQLASTSTSFSLMOWSNILKTHFT 203
Qy 69 LTER-----EQIROGLELQKVLPGDITYMHGFERASEQIYYENROGYRTA-SVIAL 122
Db 204 PTEFKSLSPQSLVDAIVQ-----GLTYTASGIQKVKELFHSKNGARKSAKILLIYI 258
Qy 123 TDGELHEDLPFYSE--REANRSRDLAGIYVCVGD-FNE-TOLARI-----ADSKDHVF 173
Db 259 TDGQKRPDLERHVIPEAKA---GIIRYALGVGDAFREPTALQELNTIGSAPSDQHV 315
Qy 174 PVNDGFQALQGIHSLKSCIEILAEPSITCAGESFQVVRNGFRHARNVD 227
Db 316 KVGN-FVALRSIORIOEK---IFALGTEBSRSSSFQHEMSQEFSSALSMD 364

RESULT 2

US-08-485-618-46
; Sequence 46, Application US/08485618
; Patent No. 5728533
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485, 618
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286, 889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32797
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1155 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-485-618-46

Query Match 10.8%; Score 166; DB 1; Length 1155;
Best Local Similarity 28.2%; Pred. No. 6.9e-09;
Matches 66; Conservative 42; Mismatches 96; Indels 30; Gaps 12;

Qy 11 PACYG-GFDLYFIIDSGSV-LHHNNEIYFYVEQLAHKFIISPOLRMSTIVSTGTTLMK 68
Db 144 PEPGQEMDIAFLIDSGSIDQSDFTQMKDFVKALMGQLASTSTSFSLMOWSNILKTHFT 203

Qy 69 LTER-----EQIROGLELQKVLPGDITYMHGFERASEQIYYENROGYRTA-SVIAL 122
Db 204 PTEFKSLSPQSLVDAIVQ-----GLTYTASGIQKVKELFHSKNGARKSAKILLIYI 258
Qy 123 TDGELHEDLPFYSE--REANRSRDLAGIYVCVGD-FNE-TOLARI-----ADSKDHVF 173
Db 259 TDGQKRPDLERHVIPEAKA---GIIRYALGVGDAFREPTALQELNTIGSAPSDQHV 315
Qy 174 PVNDGFQALQGIHSLKSCIEILAEPSITCAGESFQVVRNGFRHARNVD 227
Db 316 KVGN-FVALRSIORIOEK---IFALGTEBSRSSSFQHEMSQEFSSALSMD 364

RESULT 3

US-08-362-652-46
; Sequence 46, Application US/08362652
; Patent No. 5766850
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,652
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286, 889
; FILING DATE: 5-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32391
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1155 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-362-652-46

Query Match 10.8%; Score 166; DB 1; Length 1155;
Best Local Similarity 28.2%; Pred. No. 6.9e-09;
Matches 66; Conservative 42; Mismatches 96; Indels 30; Gaps 12;

Qy 11 PACYG-GFDLYFIIDSGSV-LHHNNEIYFYVEQLAHKFIISPOLRMSTIVSTGTTLMK 68
Db 144 PEPGQEMDIAFLIDSGSIDQSDFTQMKDFVKALMGQLASTSTSFSLMOWSNILKTHFT 203

Qy 69 LTER-----EQIROGLELQKVLPGDITYMHGFERASEQIYYENROGYRTA-SVIAL 122
Db 204 PTEFKSLSPQSLVDAIVQ-----GLTYTASGIQKVKELFHSKNGARKSAKILLIYI 258
Qy 123 TDGELHEDLPFYSE--REANRSRDLAGIYVCVGD-FNE-TOLARI-----ADSKDHVF 173

Db 259 TDGKFRDPLEYRHHVIPAERA---GIIRYAIQVGDAREPTALQELNTTIGSAPSQDHVF 315
Qy 174 PVNDGFOALQGIHSLKSCIEILAAEPSTICAGESFOVVVNGNGFRHARNVD 227
Db 316 KVGK-FVALRSTIQIOIEK----IFAIEGTESRSSSFQHEMSQEGFSALMSMD 364

RESULT 4

US-08-605-672-46
; Sequence 46, Application US/08605672
; Patent No. 5817515
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vlieten, Monica
; TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/605.672
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1155 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-605-672-46

Query Match 10.8%; Score 166; DB 2; Length 1155;

Best Local Similarity 28.2%; Pred. No. 6.9e-09;

Matches 66; Conservative 42; Mismatches 96; Indels 30; Gaps 12;

Qy 11 PACYG-GEFDYFLIDKSGSV-LHHMNEIYFVEQLAHKFSIPOLRMSFIVSTGTITMK 68
Db 144 PEGCGQEMDLAFLIDGSGISDQSPFTQMKDFVAKLMGQLASTSTSFSLMQISNLKTHFT 203
Qy 69 LTEDR-----EQIRQGLEBLQKVLPGGDTYMHGFEFASQIYYENRQGYRTA-SVIAL 122
Db 204 FTEFKSLSPOSIVDAIVQLQ-----GLTYTASGIQKVVKELEPHSKNGARKSAKKIILIVI 258
Qy 123 TDGELHEDLFYYS--REARSRDLGAIYCVGVKD-FNE-TOLARI-----ADSKDHVF 173
Db 259 TDGKFRDPLEYRHHVIPAERA---GIIRYAIQVGDAREPTALQELNTTIGSAPSQDHVF 315

Qy 174 PVNDGFOALQGIHSLKSCIEILAAEPSTICAGESFOVVVNGNGFRHARNVD 227
Db 316 KVGK-FVALRSTIQIOIEK----IFAIEGTESRSSSFQHEMSQEGFSALMSMD 364

RESULT 5

US-08-482-293A-46
; Sequence 46, Application US/08482293A
; Patent No. 5831029
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vlieten, Monica
; TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482.293A
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1155 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-482-293A-46

Query Match 10.8%; Score 166; DB 2; Length 1155;

Best Local Similarity 28.2%; Pred. No. 6.9e-09;

Matches 66; Conservative 42; Mismatches 96; Indels 30; Gaps 12;

Qy 11 PACYG-GEFDYFLIDKSGSV-LHHMNEIYFVEQLAHKFSIPOLRMSFIVSTGTITMK 68
Db 144 PEGCGQEMDLAFLIDGSGISDQSPFTQMKDFVAKLMGQLASTSTSFSLMQISNLKTHFT 203
Qy 69 LTEDR-----EQIRQGLEBLQKVLPGGDTYMHGFEFASQIYYENRQGYRTA-SVIAL 122
Db 204 FTEFKSLSPOSIVDAIVQLQ-----GLTYTASGIQKVVKELEPHSKNGARKSAKKIILIVI 258
Qy 123 TDGELHEDLFYYS--REARSRDLGAIYCVGVKD-FNE-TOLARI-----ADSKDHVF 173
Db 259 TDGKFRDPLEYRHHVIPAERA---GIIRYAIQVGDAREPTALQELNTTIGSAPSQDHVF 315
Qy 174 PVNDGFOALQGIHSLKSCIEILAAEPSTICAGESFOVVVNGNGFRHARNVD 227

Db 316 KVGN-FVALRSIQROIQEK----IFAIEGTSSRSSSFQHEMSQEGFSSALSMD 364

```
RESULT 6
US-08-943-363-46
; Sequence 46, Application US/08943363
; Patent No. 5837478
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5837478el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Seear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943.363
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173.497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286.889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362.652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: William Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1155 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-943-363-46
```

```
Query Match 10.8%; Score 166; DB 2; Length 1155;
Best Local Similarity 28.2%; Pred. No. 6.9e-09;
Matches 66; Conservative 42; Mismatches 96; Indels 30; Gaps 12;

Db 11 PACVY-GFDLYFIIDKSGSV-LHHNNEIYVFEQLAHKFIISPOLRMSFIVFSTGTTLMK 68
144 PEPCQGMDFALFDGSGSIDQSDFTQMKDFVAKLMQGLASTSTSFSLMQSNIKLKHTFT 203
69 LTEDR-----EQIRQGLELQKVLPGGDTYMHGFERASEQIYENRQGYRTA-SVIAL 122
204 FTEFKSSLSPOSIVDAIVQLQ-----GLTYTASGIQKVVLELPHSKNGAKRSACKILIVI 258
123 TDGELHEDLFFYSE--REANRSRDLAGAIVYCVKQD-FNE-TQLART-----ADSKDHF 173
259 TDGQKFRDPLEYRHVIVPEAKA---GIIRYAIGVDAREPTALQELNTLTIGSAPSODHVF 315
174 PVNDGFQALQGIHSLKSCIEILAAPEPTICAGSFQVAVVNGGFRHARNVD 227
316 KVGN-FVALRSIQROIQEK----IFAIEGTSSRSSSFQHEMSQEGFSSALSMD 364
```

```
RESULT 7
US-09-193-043-46
; Sequence 46, Application US/09193043
; Patent No. 6251395
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6251395el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/193.043
; CURRENT FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173.497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286.889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362.652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943.363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 46
; LENGTH: 1155
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-193-043-46
```

```
Query Match 10.8%; Score 166; DB 3; Length 1155;
Best Local Similarity 28.2%; Pred. No. 6.9e-09;
Matches 66; Conservative 42; Mismatches 96; Indels 30; Gaps 12;

Db 11 PACVY-GFDLYFIIDKSGSV-LHHNNEIYVFEQLAHKFIISPOLRMSFIVFSTGTTLMK 68
144 PEPCQGMDFALFDGSGSIDQSDFTQMKDFVAKLMQGLASTSTSFSLMQSNIKLKHTFT 203
69 LTEDR-----EQIRQGLELQKVLPGGDTYMHGFERASEQIYENRQGYRTA-SVIAL 122
204 FTEFKSSLSPOSIVDAIVQLQ-----GLTYTASGIQKVVLELPHSKNGAKRSACKILIVI 258
123 TDGELHEDLFFYSE--REANRSRDLAGAIVYCVKQD-FNE-TQLART-----ADSKDHF 173
259 TDGQKFRDPLEYRHVIVPEAKA---GIIRYAIGVDAREPTALQELNTLTIGSAPSODHVF 315
174 PVNDGFQALQGIHSLKSCIEILAAPEPTICAGSFQVAVVNGGFRHARNVD 227
316 KVGN-FVALRSIQROIQEK----IFAIEGTSSRSSSFQHEMSQEGFSSALSMD 364

RESULT 8
US-09-688-307A-46
; Sequence 46, Application US/09688307A
; Patent No. 6432404
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6432404el Human Beta-2
; FILE REFERENCE: 27866/36646
; CURRENT APPLICATION NUMBER: US/09/688.307A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 09/193.043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/605.672
; PRIOR FILING DATE: 1996-02-22
; PRIOR APPLICATION NUMBER: 08/173.497
; PRIOR FILING DATE: 1993-12-23
; PRIOR APPLICATION NUMBER: 08/286.889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362.652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943.363
; PRIOR FILING DATE: 1997-10-03
```

NUMBER OF SEQ ID NOS: 114
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO: 46
LENGTH: 1155
TYPE: PRT
ORGANISM: Mus musculus
US-09-688-307A-46

Query Match 10.8%; Score 166; DB 4; Length 1155;
Best Local Similarity 28.2%; Pred. No. 6.9e-09;
Matches 66; Conservative 42; Mismatches 96; Indels 30; Gaps 12;

QY 11 PACG-GFDLYFIIDKSGSV-LHHMNEIYFVEQLAHKFIISPOLRMSPIVSTGTITLMK 68
DB 144 PECGQGMIDIAFLIDSGSISDQSDFTQMKDFVAKLMQGLASTSTSPSLMQSINLTKHTFT 203
QY 69 LTER-----EQIRQGLEELQKVLPGSDTYMHGFERASEQIYYENRQGYTA-SVIAL 122
DB 204 FTEKSSLSPOSIVDAIVQLQ-----GLTYASGIQKVKELFHSKNGAKSAKKILIVI 258
QY 123 TDGELHEDLFYYSR--REANRSRLGAIYVCVGD-FNE-TOLARI-----ADSKDHVF 173
DB 259 TDGQKFRDPLEYRHVIPAERA---GIIRVAIGVGDAREPTALQELNTIGSAPSDHVF 315
QY 174 PVNDGFQALQGIHSILKSCIEILAEPTTCAGESFQVVVNGNFRHARNVD 227
DB 316 KVGN-FVALRSIQRIQIOEK---IFAIEGTESRSSSFQHEMSQEGFSALSMD 364

RESULT 9

US-09-350-259-46
Sequence 46, Application US/09350259
Patent No. 6620915
GENERAL INFORMATION:
APPLICANT: Gallatin, Michael W.
TITLE OF INVENTION: No. 6620915el Human 2
FILE REFERENCE: 27866/35004
CURRENT APPLICATION NUMBER: US/09/350.259
EARLIER FILING DATE: 1999-07-08
EARLIER APPLICATION NUMBER: 09/193.043
EARLIER FILING DATE: 1998-11-16
EARLIER APPLICATION NUMBER: 08/173.497
EARLIER FILING DATE: 1993-12-23
EARLIER APPLICATION NUMBER: 08/286.889
EARLIER FILING DATE: 1994-06-05
EARLIER APPLICATION NUMBER: 08/362.652
EARLIER FILING DATE: 1994-12-21
EARLIER APPLICATION NUMBER: 08/943.363
EARLIER FILING DATE: 1997-10-03
NUMBER OF SEQ ID NOS: 114
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO: 46
LENGTH: 1155
TYPE: PRT
ORGANISM: Mus musculus
US-09-350-259-46

Query Match 10.8%; Score 166; DB 4; Length 1155;
Best Local Similarity 28.2%; Pred. No. 6.9e-09;

Matches 66; Conservative 42; Mismatches 96; Indels 30; Gaps 12;

QY 11 PACG-GFDLYFIIDKSGSV-LHHMNEIYFVEQLAHKFIISPOLRMSPIVSTGTITLMK 68
DB 144 PECGQGMIDIAFLIDSGSISDQSDFTQMKDFVAKLMQGLASTSTSPSLMQSINLTKHTFT 203
QY 69 LTER-----EQIRQGLEELQKVLPGSDTYMHGFERASEQIYYENRQGYTA-SVIAL 122
DB 204 FTEKSSLSPOSIVDAIVQLQ-----GLTYASGIQKVKELFHSKNGAKSAKKILIVI 258
QY 123 TDGELHEDLFYYSR--REANRSRLGAIYVCVGD-FNE-TOLARI-----ADSKDHVF 173
DB 259 TDGQKFRDPLEYRHVIPAERA---GIIRVAIGVGDAREPTALQELNTIGSAPSDHVF 315

QY 174 PVNDGFQALQGIHSILKSCIEILAEPTTCAGESFQVVVNGNFRHARNVD 227
DB 316 KVGN-FVALRSIQRIQIOEK---IFAIEGTESRSSSFQHEMSQEGFSALSMD 364

RESULT 10

US-08-485-618-53
Sequence 53, Application US/08485618
Patent No. 5728533
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vlieten, Monica
TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Seair Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485.618
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32797
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 1161 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-485-618-53

Query Match 10.8%; Score 166; DB 1; Length 1161;
Best Local Similarity 28.2%; Pred. No. 6.9e-09;
Matches 66; Conservative 42; Mismatches 96; Indels 30; Gaps 12;

QY 11 PACG-GFDLYFIIDKSGSV-LHHMNEIYFVEQLAHKFIISPOLRMSPIVSTGTITLMK 68
DB 144 PECGQGMIDIAFLIDSGSISDQSDFTQMKDFVAKLMQGLASTSTSPSLMQSINLTKHTFT 203
QY 69 LTER-----EQIRQGLEELQKVLPGSDTYMHGFERASEQIYYENRQGYTA-SVIAL 122
DB 204 FTEKSSLSPOSIVDAIVQLQ-----GLTYASGIQKVKELFHSKNGAKSAKKILIVI 258
QY 123 TDGELHEDLFYYSR--REANRSRLGAIYVCVGD-FNE-TOLARI-----ADSKDHVF 173
DB 259 TDGQKFRDPLEYRHVIPAERA---GIIRVAIGVGDAREPTALQELNTIGSAPSDHVF 315
QY 174 PVNDGFQALQGIHSILKSCIEILAEPTTCAGESFQVVVNGNFRHARNVD 227

Db 316 KVG-N-FVALRSIQRIQIEK---IFAIEGTSSRSSSFQHEMSQEGFSSALSM 364

RESULT 11

US-08-362-652-53
; Sequence 53, Application US/08362652
; Patent No. 5766850
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,652
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32391
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1161 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-362-652-53

Query Match 10.8%; Score 166; DB 1; Length 1161;

Best Local Similarity 28.2%; Pred. No. 6.9e-09;

Matches 66; Conservative 42; Mismatches 96; Indels 30; Gaps 12;

QY 11 PACYG-GFDLYFLIDSGSV-LHHMNEIYFVEQLAHKFISPOLRMSFIYFSTRGTLTK 68
Db 144 PECPGQEMDIADFLIDSGSIDSDPTQKDFKALMGQLASTSFSLMOWSNILKTHFT 203
QY 69 LTEDR-----EQIRGLEBELQKVLPGDVTYMHGEPERASEQIYENROGYTA-SVIAL 122
Db 204 FTEFKSLSPQSLVDALVQLQ-----GLTYTAGSIQKVVELFHSKNGARKSAKKILIVI 258
QY 123 TDGELHEDLFYSE--REANRSRLGAIYVCVVKD-FNE-TOLARI-----ADSKDHVF 173
Db 259 TDGQKFRDPLEYRHVIPAELKA---GIRYALGVGDAFREPTALQELINTIGSAPSODHVF 315
QY 174 PVNDGFQALQGIHLSILKSKCIRILAAPSTICAGESFOVVVANGNFRHARNVD 227
Db 316 KVG-N-FVALRSIQRIQIEK---IFAIEGTSSRSSSFQHEMSQEGFSSALSM 364

RESULT 12

US-08-605-672-53
; Sequence 53, Application US/08605672
; Patent No. 5817515
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/605,672
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32604
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1161 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-605-672-53

Query Match 10.8%; Score 166; DB 2; Length 1161;

Best Local Similarity 28.2%; Pred. No. 6.9e-09;

Matches 66; Conservative 42; Mismatches 96; Indels 30; Gaps 12;

QY 11 PACYG-GFDLYFLIDSGSV-LHHMNEIYFVEQLAHKFISPOLRMSFIYFSTRGTLTK 68
Db 144 PECPGQEMDIADFLIDSGSIDSDPTQKDFKALMGQLASTSFSLMOWSNILKTHFT 203
QY 69 LTEDR-----EQIRGLEBELQKVLPGDVTYMHGEPERASEQIYENROGYTA-SVIAL 122
Db 204 FTEFKSLSPQSLVDALVQLQ-----GLTYTAGSIQKVVELFHSKNGARKSAKKILIVI 258
QY 123 TDGELHEDLFYSE--REANRSRLGAIYVCVVKD-FNE-TOLARI-----ADSKDHVF 173
Db 259 TDGQKFRDPLEYRHVIPAELKA---GIRYALGVGDAFREPTALQELINTIGSAPSODHVF 315
QY 174 PVNDGFQALQGIHLSILKSKCIRILAAPSTICAGESFOVVVANGNFRHARNVD 227
Db 316 KVG-N-FVALRSIQRIQIEK---IFAIEGTSSRSSSFQHEMSQEGFSSALSM 364

RESULT 13

US-08-482-293A-53
; Sequence 53, Application US/08482293A


```
Patent No. 5831029
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vlieten, Monica
TITLE OF INVENTION: No. 5831029e1 Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,293A
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 1161 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-482-293A-53

Query Match 10.8%; Score 166; DB 2; Length 1161;
Best Local Similarity 28.2%; Pred. No. 6.9e-09;
Matches 66; Conservative 42; Mismatches 96; Indels 30; Gaps 12;

QY 11 PACYG-GFDLYFILDKSGSV-LHNMNEIYFVEQLAHKFTSPQLRMGFIVSTRTGLMK 68
DB 144 PECGQGMDFLIDSGSIDQSPFTQMKDFVKALMGQLASTSTFSIMQYSNILKTHFT 203
QY 69 LTEDR-----EQIRQGLELQKVLPGSDTYMHGFEFASBOIYYENRQGYRTA-SVIAL 122
DB 204 FTEFKSLSPQSLVDALVQLQ-----GLTYASGIQKVKELPHSKNGARKSAKKILIVI 258
QY 123 TDGELHDLFFYSR--REANRSRLGAIYVCVYKD-FNE-TQLARI-----ADSKDHF 173
DB 259 TDGQKFRDPLEYRHVYIPEAKA---GIIRYAIQVDAFREPTALQELNTIGSAPSQDHF 315
QY 174 PVNDGFOALOGIHSILKSCIEILAEPTICAGSEFOVVVNGNFRHARNVD 227
DB 316 KVGN-FVALRSIQROIERK---IFALEGTESRSSSFQHEMSQEGFSSALSM 364

RESULT 14
US-08-943-363-53
Sequence 53, Application US/08943363
Patent No. 5837478
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vlieten, Monica
```

```
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vlieten, Monica
TITLE OF INVENTION: No. 5837478e1 Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,363
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 1161 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-943-363-53

Query Match 10.8%; Score 166; DB 2; Length 1161;
Best Local Similarity 28.2%; Pred. No. 6.9e-09;
Matches 66; Conservative 42; Mismatches 96; Indels 30; Gaps 12;

QY 11 PACYG-GFDLYFILDKSGSV-LHNMNEIYFVEQLAHKFTSPQLRMGFIVSTRTGLMK 68
DB 144 PECGQGMDFLIDSGSIDQSPFTQMKDFVKALMGQLASTSTFSIMQYSNILKTHFT 203
QY 69 LTEDR-----EQIRQGLELQKVLPGSDTYMHGFEFASBOIYYENRQGYRTA-SVIAL 122
DB 204 FTEFKSLSPQSLVDALVQLQ-----GLTYASGIQKVKELPHSKNGARKSAKKILIVI 258
QY 123 TDGELHDLFFYSR--REANRSRLGAIYVCVYKD-FNE-TQLARI-----ADSKDHF 173
DB 259 TDGQKFRDPLEYRHVYIPEAKA---GIIRYAIQVDAFREPTALQELNTIGSAPSQDHF 315
QY 174 PVNDGFOALOGIHSILKSCIEILAEPTICAGSEFOVVVNGNFRHARNVD 227
DB 316 KVGN-FVALRSIQROIERK---IFALEGTESRSSSFQHEMSQEGFSSALSM 364

RESULT 15
US-09-193-043-53
Sequence 53, Application US/09193043
Patent No. 6251395
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vlieten, Monica
```

TITLE OF INVENTION: No. 6251395e1 Human 2
FILE REFERENCE: 27866/35004
CURRENT APPLICATION NUMBER: US/09/193,043
CURRENT FILING DATE: 1998-11-16
EARLIER APPLICATION NUMBER: 08/173,497
EARLIER FILING DATE: 1993-12-23
EARLIER APPLICATION NUMBER: 08/286,889
EARLIER FILING DATE: 1994-08-05
EARLIER APPLICATION NUMBER: 08/362,652
EARLIER FILING DATE: 1994-12-21
EARLIER APPLICATION NUMBER: 08/943,363
EARLIER FILING DATE: 1997-10-03
NUMBER OF SEQ ID NOS: 114
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 53
LENGTH: 1161
TYPE: PRT
ORGANISM: Mus musculus
US-09-193-043-53

Query Match 10.8%; Score 166; DB 3; Length 1161;

Best Local Similarity 28.2%; Pred. No. 6.9e-09;
Matches 66; Conservative 42; Mismatches 96; Indels 30; Gaps 12;

QY 11 PACYG-GFDLYFLIDKSGSV-LHMNEIYFVEQLAKKFIPOLRMSFIVFSTRTTLMK 68
DB 144 PECFQEMDIAPFLIDSGSIDOSDFTQMKDFKALMGQLASTSTSFSLMOYSNLIKTHPT 203
QY 69 LTEDR-----EQIRQGLELQKVLGSDTYMHGEPERASBOIYYENRQGYRTA-SVITIAL 122
DB 204 FTERKSSLSPOSVDALVQLQ-----GLTYTASGIQKVVKELFHSKNGARKSAKKILIVI 258
QY 123 TDGELHEDLFFYSE--REANRSRDIGALVYCVGVKD-FNE-TOLARI-----ADSKDHF 173
DB 259 TDGQKFRDPLEYRHYIPEAKA---GIIRVAIGVDARFEPALQELNTIGSAPSODHVF 315
QY 174 PVNDGFQALQGIHSILKSCIEILAAEPSTICAGESFOVVYVRNGFRHARNVD 227
DB 316 KVGN-FVALRSIORIOIEK---IFAIEGTSSRSSSFQHEMSQEGFSSALSMD 364

Search completed: June 13, 2005, 20:05:07
Job time : 29.681 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 20:03:29 ; Search time 94.5881 Seconds
(without alignments)
1195.537 Million cell updates/sec

Title: US-09-970-076-2_COPY_27_321
Perfect score: 1536
Sequence: 1 GGGGRREDGGPACGFGFDLY.....GLSFSSVITTTTCHSDGS 295

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1710399 seqs, 38334425 residues

Total number of hits satisfying chosen parameters: 1710399

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications_AA:*
1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pcp:*
2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pcp:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pcp:*
4: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pcp:*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pcp:*
6: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pcp:*
7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pcp:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pcp:*
9: /cgn2_6/ptodata/1/pubppaa/US09A_PUBCOMB.pcp:*
10: /cgn2_6/ptodata/1/pubppaa/US09B_PUBCOMB.pcp:*
11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pcp:*
12: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pcp:*
13: /cgn2_6/ptodata/1/pubppaa/US10A_PUBCOMB.pcp:*
14: /cgn2_6/ptodata/1/pubppaa/US10B_PUBCOMB.pcp:*
15: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pcp:*
16: /cgn2_6/ptodata/1/pubppaa/US10D_PUBCOMB.pcp:*
17: /cgn2_6/ptodata/1/pubppaa/US10E_PUBCOMB.pcp:*
18: /cgn2_6/ptodata/1/pubppaa/US10F_NEW_PUB.pcp:*
19: /cgn2_6/ptodata/1/pubppaa/US11A_PUBCOMB.pcp:*
20: /cgn2_6/ptodata/1/pubppaa/US11_NEW_PUB.pcp:*
21: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pcp:*
22: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pcp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1536	100.0	403	11 US-09-833-245-621	Sequence 621, App
2	1536	100.0	564	10 US-09-918-715-187	Sequence 187, App
3	1536	100.0	564	10 US-09-918-715-232	Sequence 232, App
4	1536	100.0	564	14 US-10-301-822-199	Sequence 199, App
5	1536	100.0	564	16 US-10-408-765A-1823	Sequence 1823, App
6	1536	100.0	564	16 US-10-474-794-187	Sequence 187, App
7	1536	100.0	564	16 US-10-474-794-232	Sequence 232, App
8	1532	99.7	551	14 US-10-038-307-18	Sequence 18, App1
9	1532	99.7	551	14 US-10-201-292-18	Sequence 18, App1
10	1524	99.2	403	11 US-09-833-245-620	Sequence 620, App1
11	1520	99.0	333	10 US-09-796-753-12	Sequence 12, App1

12	1520	99.0	333	14 US-10-038-307-2	Sequence 2, App1
13	1520	99.0	333	14 US-10-201-292-2	Sequence 2, App1
14	1520	99.0	345	14 US-10-038-307-24	Sequence 24, App1
15	1520	99.0	345	14 US-10-201-292-24	Sequence 24, App1
16	1520	99.0	564	14 US-10-038-307-20	Sequence 20, App1
17	1520	99.0	564	14 US-10-201-292-20	Sequence 20, App1
18	1519	98.9	562	10 US-09-918-715-194	Sequence 194, App
19	1519	98.9	562	10 US-09-918-715-301	Sequence 301, App
20	1519	98.9	562	16 US-10-474-794-194	Sequence 194, App
21	1519	98.9	562	16 US-10-474-794-301	Sequence 301, App
22	1507	98.1	328	14 US-10-038-307-26	Sequence 26, App1
23	1507	98.1	328	14 US-10-201-292-26	Sequence 26, App1
24	1505.5	98.0	342	14 US-10-038-307-22	Sequence 22, App1
25	1505.5	98.0	342	14 US-10-201-292-22	Sequence 22, App1
26	1500	97.7	543	14 US-10-038-307-14	Sequence 14, App1
27	1500	97.7	543	14 US-10-038-307-16	Sequence 16, App1
28	1500	97.7	543	14 US-10-201-292-16	Sequence 16, App1
29	1500	97.7	543	14 US-10-201-292-14	Sequence 14, App1
30	1490	97.0	543	14 US-10-038-307-10	Sequence 10, App1
31	1490	97.0	543	14 US-10-201-292-10	Sequence 10, App1
32	1479	96.3	534	14 US-10-038-307-12	Sequence 12, App1
33	1479	96.3	534	14 US-10-201-292-12	Sequence 12, App1
34	1419	92.4	529	14 US-10-201-292-36	Sequence 36, App1
35	1305.5	85.0	504	14 US-10-201-292-36	Sequence 36, App1
36	1178	76.7	479	14 US-10-201-292-32	Sequence 32, App1
37	1064	69.3	460	14 US-10-201-292-28	Sequence 28, App1
38	1054	68.6	460	14 US-10-201-292-30	Sequence 30, App1
39	903	58.8	538	13 US-10-047-542-99	Sequence 99, App1
40	790.5	51.5	488	10 US-09-796-753-52	Sequence 52, App1
41	790.5	51.5	488	10 US-10-038-307-6	Sequence 6, App1
42	790.5	51.5	488	14 US-10-201-292-6	Sequence 6, App1
43	790.5	51.5	488	14 US-10-368-087-16	Sequence 16, App1
44	777.5	50.6	587	9 US-10-104-047-2639	Sequence 2639, App
45				US-09-764-870-312	Sequence 312, App

ALIGNMENTS

RESULT 1
US-09-833-245-621 Application US/09833245
Sequence 621, Application US/09833245
Publication No. US20040010134A1
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: PFS46PCT
CURRENT APPLICATION NUMBER: US/09/833,245
CURRENT FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 60/229,358
PRIOR FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: 60/256,931
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/199,384
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 2267
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 621
LENGTH: 403
TYPE: PRT
ORGANISM: Homo sapiens
US-09-833-245-621

Query Match 100.0%; Score 1536; DB 11; Length 403;
Best Local Similarity 100.0%; Pred. No. 7.4e-149; Indels 0; Gaps 0;
Matches 295; Conservative 0; Mismatches 0;
QY 1 GGGGRREDGGPACGFGFDLYFIIDKSGSVLHHNNEIYFYEQLAHKFISPOLHMSFVFS 60
DB 27 GGGGRREDGGPACGFGFDLYFIIDKSGSVLHHNNEIYFYEQLAHKFISPOLHMSFVFS 86
QY 61 TRGTTMLKLTEDDEQIRGIGLELQKYLPGGDIYMHGFERASQIYYENRQYRTASVII 120

Db 87 TRGTTLMKLTEDRQIRQGLELQKVLPGDVTYMHGFERASEQIYYENRGYRTASVIT 146
Qy 121 ALTDGELHEDLFFYSEREARNSRDIGAIYVCVGDVFENETOLARIADSKDHVPVNDGFQ 180
Db 147 ALTDGELHEDLFFYSEREARNSRDIGAIYVCVGDVFENETOLARIADSKDHVPVNDGFQ 206
Qy 181 ALQGIHSILKKSCEIILAAEPSTTCAGESFOVVVRGNGFRHARNDRVLCSPKINDSVT 240
Db 207 ALQGIHSILKKSCEIILAAEPSTTCAGESFOVVVRGNGFRHARNDRVLCSPKINDSVT 266
Qy 241 LNEKPSVEDVTYLLCPAPILKEVGKAKALQVSNMDGLSFSSSVIITTHCSDGS 295
Db 267 LNEKPSVEDVTYLLCPAPILKEVGKAKALQVSNMDGLSFSSSVIITTHCSDGS 321

RESULT 2

US-09-918-715-187
Sequence 187, Application US/09918715
Publication No. US20030017157A1
GENERAL INFORMATION:
APPLICANT: Brad St. Croix
APPLICANT: Bert Vogelstein
APPLICANT: Kenneth Kinzler
TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS
FILE REFERENCE: 1107.00134
CURRENT APPLICATION NUMBER: US/09/918.715
CURRENT FILING DATE: 2001-08-01
PRIOR APPLICATION NUMBER: 60/222,599
PRIOR FILING DATE: 2000-08-02
PRIOR APPLICATION NUMBER: 60/224,360
PRIOR FILING DATE: 2000-08-11
PRIOR APPLICATION NUMBER: 60/282,850
PRIOR FILING DATE: 2000-04-11
NUMBER OF SEQ ID NOS: 358
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 187
LENGTH: 564
TYPE: PRT
ALQGIHSILKKSCEIILAAEPSTTCAGESFOVVVRGNGFRHARNDRVLCSPKINDSVT
ORGANISM: Homo sapiens
US-09-918-715-187

Query Match 100.0%; Score 1536; DB 10; Length 564;
Best Local Similarity 100.0%; Pred. No. 1.2e-148;
Matches 295; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGRRRBDGGPACGCGFDLYFILDKSGSVLHHNNEIYFYEQLAHKFISPOLRMSFTVFS 60
Db 27 GGGRRRBDGGPACGCGFDLYFILDKSGSVLHHNNEIYFYEQLAHKFISPOLRMSFTVFS 86
Qy 61 TRGTTLMKLTEDRQIRQGLELQKVLPGDVTYMHGFERASEQIYYENRGYRTASVIT 120
Db 87 TRGTTLMKLTEDRQIRQGLELQKVLPGDVTYMHGFERASEQIYYENRGYRTASVIT 146
Qy 121 ALTDGELHEDLFFYSEREARNSRDIGAIYVCVGDVFENETOLARIADSKDHVPVNDGFQ 180
Db 147 ALTDGELHEDLFFYSEREARNSRDIGAIYVCVGDVFENETOLARIADSKDHVPVNDGFQ 206
Qy 181 ALQGIHSILKKSCEIILAAEPSTTCAGESFOVVVRGNGFRHARNDRVLCSPKINDSVT 240
Db 207 ALQGIHSILKKSCEIILAAEPSTTCAGESFOVVVRGNGFRHARNDRVLCSPKINDSVT 266
Qy 241 LNEKPSVEDVTYLLCPAPILKEVGKAKALQVSNMDGLSFSSSVIITTHCSDGS 295
Db 267 LNEKPSVEDVTYLLCPAPILKEVGKAKALQVSNMDGLSFSSSVIITTHCSDGS 321

RESULT 3

US-09-918-715-232
Sequence 232, Application US/09918715
Publication No. US20030017157A1
GENERAL INFORMATION:
APPLICANT: Brad St. Croix
APPLICANT: Bert Vogelstein

APPLICANT: Kenneth Kinzler
TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS
FILE REFERENCE: 1107.00134
CURRENT APPLICATION NUMBER: US/09/918.715
CURRENT FILING DATE: 2001-08-01
PRIOR APPLICATION NUMBER: 60/222,599
PRIOR FILING DATE: 2000-08-02
PRIOR APPLICATION NUMBER: 60/224,360
PRIOR FILING DATE: 2000-08-11
PRIOR APPLICATION NUMBER: 60/282,850
PRIOR FILING DATE: 2000-04-11
NUMBER OF SEQ ID NOS: 358
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 232
LENGTH: 564
TYPE: PRT
ORGANISM: Homo sapiens
US-09-918-715-232

Query Match 100.0%; Score 1536; DB 10; Length 564;
Best Local Similarity 100.0%; Pred. No. 1.2e-148;
Matches 295; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGRRRBDGGPACGCGFDLYFILDKSGSVLHHNNEIYFYEQLAHKFISPOLRMSFTVFS 60
Db 27 GGGRRRBDGGPACGCGFDLYFILDKSGSVLHHNNEIYFYEQLAHKFISPOLRMSFTVFS 86
Qy 61 TRGTTLMKLTEDRQIRQGLELQKVLPGDVTYMHGFERASEQIYYENRGYRTASVIT 120
Db 87 TRGTTLMKLTEDRQIRQGLELQKVLPGDVTYMHGFERASEQIYYENRGYRTASVIT 146
Qy 121 ALTDGELHEDLFFYSEREARNSRDIGAIYVCVGDVFENETOLARIADSKDHVPVNDGFQ 180
Db 147 ALTDGELHEDLFFYSEREARNSRDIGAIYVCVGDVFENETOLARIADSKDHVPVNDGFQ 206
Qy 181 ALQGIHSILKKSCEIILAAEPSTTCAGESFOVVVRGNGFRHARNDRVLCSPKINDSVT 240
Db 207 ALQGIHSILKKSCEIILAAEPSTTCAGESFOVVVRGNGFRHARNDRVLCSPKINDSVT 266
Qy 241 LNEKPSVEDVTYLLCPAPILKEVGKAKALQVSNMDGLSFSSSVIITTHCSDGS 295
Db 267 LNEKPSVEDVTYLLCPAPILKEVGKAKALQVSNMDGLSFSSSVIITTHCSDGS 321

RESULT 4

US-10-301-822-199
Sequence 199, Application US/10301822
Publication No. US20030148410A1
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc.
APPLICANT: Berger, Allison
APPLICANT: Guillemette, Tracy L.
APPLICANT: Kametkar, Shubhangi
APPLICANT: Schlegel, Robert
APPLICANT: Monahan, John E.
APPLICANT: Thibodeau, Stephen N.
APPLICANT: Burgart, Lawrence J.
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
TITLE OF INVENTION: THERAPY OF COLON CANCER
FILE REFERENCE: MEM01-029P2RNM
CURRENT APPLICATION NUMBER: US/10/301.822
CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 60/339,971
PRIOR FILING DATE: 2001-12-10
PRIOR APPLICATION NUMBER: US 60/361,978
PRIOR FILING DATE: 2002-03-05
PRIOR APPLICATION NUMBER: US 60/381,988
PRIOR FILING DATE: 2002-05-20
NUMBER OF SEQ ID NOS: 228
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 199
LENGTH: 564

```
/ TYPE: PRT
/ ORGANISM: Homo Sapiens
US-10-301-822-199

Query Match
Best Local Similarity 100.0%; Score 1536; DB 14; Length 564;
Matches 295; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGRRREDGPGACGCGPDLVFIIDKSGSVLHNMNIYFVEQLAHKFIISPOLRMSFIYFS 60
DB 27 GGGRRREDGPGACGCGPDLVFIIDKSGSVLHNMNIYFVEQLAHKFIISPOLRMSFIYFS 86
QY 61 TRGTTLMKLTEDRQIRQGLELOKVLPGDVTYHMEGERASEQIYYENRQGYRTASVYI 120
DB 87 TRGTTLMKLTEDRQIRQGLELOKVLPGDVTYHMEGERASEQIYYENRQGYRTASVYI 146
QY 121 ALTDEGLHEDLFYFSEREARNSRDGAIVYCVGVDFNETQLARIADSKDHVPVNDGFQ 180
DB 147 ALTDEGLHEDLFYFSEREARNSRDGAIVYCVGVDFNETQLARIADSKDHVPVNDGFQ 206
QY 181 ALQGIHSLILKSCIEILAAEPSTICAGESFQVVRNGGFRHARNDVRLCSFKINDSVT 240
DB 207 ALQGIHSLILKSCIEILAAEPSTICAGESFQVVRNGGFRHARNDVRLCSFKINDSVT 266
QY 241 LNEKPFSEVDYLLCPAPILKEVGKKAALQVSMNDGSLFISSSVIITTHCSDGS 295
DB 267 LNEKPFSEVDYLLCPAPILKEVGKKAALQVSMNDGSLFISSSVIITTHCSDGS 321

RESULT 5
US-10-408-765A-1823
/ Sequence 1823, Application US/10408765A
/ Publication No. US20040101874A1
/ GENERAL INFORMATION:
/ APPLICANT: Ghosh, Soumitra S.
/ APPLICANT: Fahy, Boia D.
/ APPLICANT: Zhang, Bing
/ APPLICANT: Gibson, Steven W.
/ APPLICANT: Taylor, Steven W.
/ APPLICANT: Glenn, Gary M.
/ APPLICANT: Wernock, Dale E.
/ TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
/ TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
/ FILE REFERENCE: 660088.465
/ CURRENT APPLICATION NUMBER: US/10/408, 765A
/ CURRENT FILING DATE: 2003-04-04
/ NUMBER OF SEQ ID NOS: 3077
/ SOFTWARE: FastSeq For Windows Version 4.0
/ SEQ ID NO 1823
/ LENGTH: 564
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-408-765A-1823

Query Match
Best Local Similarity 100.0%; Score 1536; DB 16; Length 564;
Matches 295; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGRRREDGPGACGCGPDLVFIIDKSGSVLHNMNIYFVEQLAHKFIISPOLRMSFIYFS 60
DB 27 GGGRRREDGPGACGCGPDLVFIIDKSGSVLHNMNIYFVEQLAHKFIISPOLRMSFIYFS 86
QY 61 TRGTTLMKLTEDRQIRQGLELOKVLPGDVTYHMEGERASEQIYYENRQGYRTASVYI 120
DB 87 TRGTTLMKLTEDRQIRQGLELOKVLPGDVTYHMEGERASEQIYYENRQGYRTASVYI 146
QY 121 ALTDEGLHEDLFYFSEREARNSRDGAIVYCVGVDFNETQLARIADSKDHVPVNDGFQ 180
DB 147 ALTDEGLHEDLFYFSEREARNSRDGAIVYCVGVDFNETQLARIADSKDHVPVNDGFQ 206
QY 181 ALQGIHSLILKSCIEILAAEPSTICAGESFQVVRNGGFRHARNDVRLCSFKINDSVT 240
DB 207 ALQGIHSLILKSCIEILAAEPSTICAGESFQVVRNGGFRHARNDVRLCSFKINDSVT 266
```

```
QY 241 LNEKPFSEVDYLLCPAPILKEVGKKAALQVSMNDGSLFISSSVIITTHCSDGS 295
DB 267 LNEKPFSEVDYLLCPAPILKEVGKKAALQVSMNDGSLFISSSVIITTHCSDGS 321

RESULT 6
US-10-474-794-187
/ Sequence 187, Application US/10474794
/ Publication No. US20040213793A1
/ GENERAL INFORMATION:
/ APPLICANT: Carson-Walter, Eleanor
/ APPLICANT: St. Croix, Brad
/ APPLICANT: Vogelstein, Bert
/ APPLICANT: Kinzler, Kenneth
/ TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS
/ FILE REFERENCE: 1107.00179
/ CURRENT APPLICATION NUMBER: US/10/474, 794
/ PRIOR FILING DATE: 2003-10-14
/ PRIOR APPLICATION NUMBER: 60/282,850
/ PRIOR FILING DATE: 2001-04-11
/ PRIOR APPLICATION NUMBER: 60/308,829
/ NUMBER OF SEQ ID NOS: 359
/ SOFTWARE: FastSeq For Windows Version 4.0
/ SEQ ID NO 187
/ LENGTH: 564
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-474-794-187

Query Match
Best Local Similarity 100.0%; Score 1536; DB 16; Length 564;
Matches 295; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGRRREDGPGACGCGPDLVFIIDKSGSVLHNMNIYFVEQLAHKFIISPOLRMSFIYFS 60
DB 27 GGGRRREDGPGACGCGPDLVFIIDKSGSVLHNMNIYFVEQLAHKFIISPOLRMSFIYFS 86
QY 61 TRGTTLMKLTEDRQIRQGLELOKVLPGDVTYHMEGERASEQIYYENRQGYRTASVYI 120
DB 87 TRGTTLMKLTEDRQIRQGLELOKVLPGDVTYHMEGERASEQIYYENRQGYRTASVYI 146
QY 121 ALTDEGLHEDLFYFSEREARNSRDGAIVYCVGVDFNETQLARIADSKDHVPVNDGFQ 180
DB 147 ALTDEGLHEDLFYFSEREARNSRDGAIVYCVGVDFNETQLARIADSKDHVPVNDGFQ 206
QY 181 ALQGIHSLILKSCIEILAAEPSTICAGESFQVVRNGGFRHARNDVRLCSFKINDSVT 240
DB 207 ALQGIHSLILKSCIEILAAEPSTICAGESFQVVRNGGFRHARNDVRLCSFKINDSVT 266
QY 241 LNEKPFSEVDYLLCPAPILKEVGKKAALQVSMNDGSLFISSSVIITTHCSDGS 295
DB 267 LNEKPFSEVDYLLCPAPILKEVGKKAALQVSMNDGSLFISSSVIITTHCSDGS 321

RESULT 7
US-10-474-794-232
/ Sequence 232, Application US/10474794
/ Publication No. US20040213793A1
/ GENERAL INFORMATION:
/ APPLICANT: Carson-Walter, Eleanor
/ APPLICANT: St. Croix, Brad
/ APPLICANT: Vogelstein, Bert
/ APPLICANT: Kinzler, Kenneth
/ TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS
/ FILE REFERENCE: 1107.00179
/ CURRENT APPLICATION NUMBER: US/10/474, 794
/ PRIOR FILING DATE: 2003-10-14
/ PRIOR APPLICATION NUMBER: 60/282,850
/ PRIOR FILING DATE: 2001-04-11
/ PRIOR APPLICATION NUMBER: 60/308,829
/ PRIOR FILING DATE: 2001-08-01
```

; NUMBER OF SEQ ID NOS: 359
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 232
 ; LENGTH: 564
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-474-794-232

Query Match 100.0%; Score 1536; DB 16; Length 564;
 Best Local Similarity 100.0%; Pred. No. 1,2e-148;
 Matches 295; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGRRRBDGGPACGCGFDLYFLIDKSGSVLHHNMEIYFVEQLAHKFIISPOLRMSFVFS 60
 DB 27 GGGRRRBDGGPACGCGFDLYFLIDKSGSVLHHNMEIYFVEQLAHKFIISPOLRMSFVFS 86
 QY 61 TRGTTLMKLTEDREIQROGLELOKVLPGSDTYMHGFEARSQIYYENRQGYRTASVII 120
 DB 87 TRGTTLMKLTEDREIQROGLELOKVLPGSDTYMHGFEARSQIYYENRQGYRTASVII 146
 QY 121 ALTDGELHEDLFFYSEREANRSRDGAIYVCVGVKDFNETQLARIADSKDHVPVNDGFQ 180
 DB 147 ALTDGELHEDLFFYSEREANRSRDGAIYVCVGVKDFNETQLARIADSKDHVPVNDGFQ 206
 QY 181 ALGGIHSILKSKCIEILAAEPSTICAGESFQVVRNGFRHARNVDRVLCSEFKINDSVT 240
 DB 207 ALGGIHSILKSKCIEILAAEPSTICAGESFQVVRNGFRHARNVDRVLCSEFKINDSVT 266
 QY 241 LNEKPSVEDTYLLCPAPILKEVGMKALQVSNMDGLSFISSVIITTHCSDG 295
 DB 267 LNEKPSVEDTYLLCPAPILKEVGMKALQVSNMDGLSFISSVIITTHCSDG 321

RESULT 8

; US-10-038-307-18
 ; Sequence 18, Application US/10038307
 ; Publication No. US20030134786A1
 ; GENERAL INFORMATION:
 ; APPLICANT: James B. ROTTMAN
 ; APPLICANT: Theresa L. O'KEEFE
 ; APPLICANT: Engin OZKANAK
 ; APPLICANT: Judith J. HEALEY
 ; TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
 ; FILE REFERENCE: 7853-253-999
 ; CURRENT FILING DATE: 2002-06-28
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 18
 ; LENGTH: 551
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-038-307-18

Query Match 99.7%; Score 1532; DB 14; Length 551;
 Best Local Similarity 100.0%; Pred. No. 3e-148;
 Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGRRRBDGGPACGCGFDLYFLIDKSGSVLHHNMEIYFVEQLAHKFIISPOLRMSFVFS 60
 DB 27 GGGRRRBDGGPACGCGFDLYFLIDKSGSVLHHNMEIYFVEQLAHKFIISPOLRMSFVFS 86
 QY 61 TRGTTLMKLTEDREIQROGLELOKVLPGSDTYMHGFEARSQIYYENRQGYRTASVII 120
 DB 87 TRGTTLMKLTEDREIQROGLELOKVLPGSDTYMHGFEARSQIYYENRQGYRTASVII 146
 QY 121 ALTDGELHEDLFFYSEREANRSRDGAIYVCVGVKDFNETQLARIADSKDHVPVNDGFQ 180
 DB 147 ALTDGELHEDLFFYSEREANRSRDGAIYVCVGVKDFNETQLARIADSKDHVPVNDGFQ 206
 QY 181 ALGGIHSILKSKCIEILAAEPSTICAGESFQVVRNGFRHARNVDRVLCSEFKINDSVT 240
 DB 207 ALGGIHSILKSKCIEILAAEPSTICAGESFQVVRNGFRHARNVDRVLCSEFKINDSVT 266

QY 241 LNEKPSVEDTYLLCPAPILKEVGMKALQVSNMDGLSFISSVIITTHCSDG 294
 DB 267 LNEKPSVEDTYLLCPAPILKEVGMKALQVSNMDGLSFISSVIITTHCSDG 320

RESULT 9

; US-10-201-292-18
 ; Sequence 18, Application US/10201292
 ; Publication No. US20030144193A1
 ; GENERAL INFORMATION:
 ; APPLICANT: James B. ROTTMAN
 ; APPLICANT: Theresa L. O'KEEFE
 ; APPLICANT: Engin OZKANAK
 ; APPLICANT: Judith J. HEALEY
 ; TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
 ; FILE REFERENCE: 7853-253-999
 ; CURRENT FILING DATE: 2003-02-14
 ; NUMBER OF SEQ ID NOS: 36
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 18
 ; LENGTH: 551
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-201-292-18

Query Match 99.7%; Score 1532; DB 14; Length 551;
 Best Local Similarity 100.0%; Pred. No. 3e-148;
 Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGRRRBDGGPACGCGFDLYFLIDKSGSVLHHNMEIYFVEQLAHKFIISPOLRMSFVFS 60
 DB 27 GGGRRRBDGGPACGCGFDLYFLIDKSGSVLHHNMEIYFVEQLAHKFIISPOLRMSFVFS 86
 QY 61 TRGTTLMKLTEDREIQROGLELOKVLPGSDTYMHGFEARSQIYYENRQGYRTASVII 120
 DB 87 TRGTTLMKLTEDREIQROGLELOKVLPGSDTYMHGFEARSQIYYENRQGYRTASVII 146
 QY 121 ALTDGELHEDLFFYSEREANRSRDGAIYVCVGVKDFNETQLARIADSKDHVPVNDGFQ 180
 DB 147 ALTDGELHEDLFFYSEREANRSRDGAIYVCVGVKDFNETQLARIADSKDHVPVNDGFQ 206
 QY 181 ALGGIHSILKSKCIEILAAEPSTICAGESFQVVRNGFRHARNVDRVLCSEFKINDSVT 240
 DB 207 ALGGIHSILKSKCIEILAAEPSTICAGESFQVVRNGFRHARNVDRVLCSEFKINDSVT 266
 QY 241 LNEKPSVEDTYLLCPAPILKEVGMKALQVSNMDGLSFISSVIITTHCSDG 294
 DB 267 LNEKPSVEDTYLLCPAPILKEVGMKALQVSNMDGLSFISSVIITTHCSDG 320

RESULT 10

; US-09-833-245-620
 ; Sequence 620, Application US/09833245
 ; Publication No. US20040010134A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Human Genome Sciences, Inc.
 ; TITLE OF INVENTION: Albumin Fusion Proteins
 ; FILE REFERENCE: p9546PCT
 ; CURRENT FILING DATE: 2001-04-12
 ; PRIOR APPLICATION NUMBER: US/09/833,245
 ; PRIOR FILING DATE: 2000-04-12
 ; PRIOR APPLICATION NUMBER: 60/229, 358
 ; PRIOR FILING DATE: 2000-04-12
 ; PRIOR APPLICATION NUMBER: 60/256, 931
 ; PRIOR FILING DATE: 2000-12-21
 ; PRIOR APPLICATION NUMBER: 60/199, 384
 ; PRIOR FILING DATE: 2000-04-25
 ; NUMBER OF SEQ ID NOS: 2267
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 620
 ; LENGTH: 403
 ; TYPE: PRT

```

/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (1175)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (1320)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (331)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (368)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-833-245-620

Query Match          99.2%; Score 1524; DB 11; Length 403;
Best Local Similarity 99.3%; Pred. No. 1,3e-147;
Matches 293; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGGRRRDGGPACGCGFDLYFLDKSGSVLHNNIEIYFVEQLAHKFTSPQLRMSFIVFS 60
Db 27 GGGRRRDGGPACGCGFDLYFLDKSGSVLHNNIEIYFVEQLAHKFTSPQLRMSFIVFS 86
Qy 61 TRGTTLMKLTEDRQIRQGLEELQKVLPGGDTYMHGFEFASQIYYENRQGYRTASVII 120
Db 87 TRGTTLMKLTEDRQIRQGLEELQKVLPGGDTYMHGFEFASQIYYENRQGYRTASVII 146
Qy 121 ALTDGELHEDLFFYSEREARSRDLGAIYCVGKDPNEFQTLARIADSKDHVPVNDGFO 180
Db 147 ALTDGELHEDLFFYSEREARSRDLGAIYCVGKDPNEFQTLARIADSKDHVPVNDGFO 206
Qy 181 ALQGIHSILKSCIEILAAEPSTTCAGSFQVYVVRNGFRRHARNDVRLCSFKINDSVT 240
Db 207 ALQGIHSILKSCIEILAAEPSTTCAGSFQVYVVRNGFRRHARNDVRLCSFKINDSVT 266
Qy 241 LNEKPSVEDTYLLCPAPILKEVGKRAALQVSNMDGLSFSSSVIITTHCSDS 295
Db 267 LNEKPSVEDTYLLCPAPILKEVGKRAALQVSNMDGLSFSSSVIITTHCSDS 321

RESULT 11
US-09-796-753-12
/ Sequence 12, Application US/09796753
/ Publication No. US20030027998A1
/ GENERAL INFORMATION:
/ APPLICANT: McCarthy, Sean A.
/ TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
/ FILE REFERENCE: 7853-227-999
/ CURRENT APPLICATION NUMBER: US/09/796,753
/ PRIOR FILING DATE: 2001-03-01
/ PRIOR APPLICATION NUMBER: 09/183,175
/ PRIOR FILING DATE: 1998-10-30
/ PRIOR APPLICATION NUMBER: 09/223,094
/ PRIOR FILING DATE: 1998-12-30
/ PRIOR APPLICATION NUMBER: 09/223,546
/ PRIOR FILING DATE: 1998-12-30
/ PRIOR APPLICATION NUMBER: 09/224,246
/ PRIOR FILING DATE: 1998-12-30
/ PRIOR APPLICATION NUMBER: 09/259,388
/ PRIOR FILING DATE: 1999-02-26
/ PRIOR APPLICATION NUMBER: 60/122,458
/ PRIOR FILING DATE: 1999-03-01
/ PRIOR APPLICATION NUMBER: 09/312,359
/ PRIOR FILING DATE: 1999-05-14
/ PRIOR APPLICATION NUMBER: 09/336,536
/ PRIOR FILING DATE: 1999-06-18
/ PRIOR APPLICATION NUMBER: 09/342,687
/ PRIOR FILING DATE: 1999-06-29
/ PRIOR APPLICATION NUMBER: 09/345,464
/ PRIOR FILING DATE: 1999-06-30
/ PRIOR APPLICATION NUMBER: 09/365,164
/ PRIOR FILING DATE: 1999-07-30
```

```

/ PRIOR APPLICATION NUMBER: 09/399,723
/ PRIOR FILING DATE: 1999-09-20
/ PRIOR APPLICATION NUMBER: 09/409,634
/ PRIOR FILING DATE: 1999-09-30
/ PRIOR APPLICATION NUMBER: 09/471,179
/ PRIOR FILING DATE: 1999-12-23
/ PRIOR APPLICATION NUMBER: 09/474,071
/ PRIOR FILING DATE: 1999-12-29
/ PRIOR APPLICATION NUMBER: 09/474,072
/ PRIOR FILING DATE: 1999-12-29
/ PRIOR APPLICATION NUMBER: 09/514,010
/ PRIOR FILING DATE: 2000-02-25
/ PRIOR APPLICATION NUMBER: 09/516,745
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: 09/572,002
/ PRIOR FILING DATE: 2000-05-14
/ PRIOR APPLICATION NUMBER: 09/597,993
/ PRIOR FILING DATE: 2000-06-19
/ PRIOR APPLICATION NUMBER: 09/599,596
/ PRIOR FILING DATE: 2000-06-22
/ PRIOR APPLICATION NUMBER: 09/630,334
/ PRIOR FILING DATE: 2000-07-31
/ PRIOR APPLICATION NUMBER: 09/606,565
/ PRIOR FILING DATE: 2000-06-29
/ PRIOR APPLICATION NUMBER: 09/606,317
/ PRIOR FILING DATE: 2000-06-29
/ PRIOR APPLICATION NUMBER: 09/665,666
/ PRIOR FILING DATE: 2000-09-20
/ PRIOR APPLICATION NUMBER: 09/677,751
/ PRIOR FILING DATE: 2000-09-30
/ NUMBER OF SEQ ID NOS: 162
/ SEQ ID NO 12
/ LENGTH: 333
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-796-753-12

Query Match          99.0%; Score 1520; DB 10; Length 333;
Best Local Similarity 100.0%; Pred. No. 2.5e-147;
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGRRRDGGPACGCGFDLYFLDKSGSVLHNNIEIYFVEQLAHKFTSPQLRMSFIVFS 60
Db 27 GGGRRRDGGPACGCGFDLYFLDKSGSVLHNNIEIYFVEQLAHKFTSPQLRMSFIVFS 86
Qy 61 TRGTTLMKLTEDRQIRQGLEELQKVLPGGDTYMHGFEFASQIYYENRQGYRTASVII 120
Db 87 TRGTTLMKLTEDRQIRQGLEELQKVLPGGDTYMHGFEFASQIYYENRQGYRTASVII 146
Qy 121 ALTDGELHEDLFFYSEREARSRDLGAIYCVGKDPNEFQTLARIADSKDHVPVNDGFO 180
Db 147 ALTDGELHEDLFFYSEREARSRDLGAIYCVGKDPNEFQTLARIADSKDHVPVNDGFO 206
Qy 181 ALQGIHSILKSCIEILAAEPSTTCAGSFQVYVVRNGFRRHARNDVRLCSFKINDSVT 240
Db 207 ALQGIHSILKSCIEILAAEPSTTCAGSFQVYVVRNGFRRHARNDVRLCSFKINDSVT 266
Qy 241 LNEKPSVEDTYLLCPAPILKEVGKRAALQVSNMDGLSFSSSVIITTHCS 292
Db 267 LNEKPSVEDTYLLCPAPILKEVGKRAALQVSNMDGLSFSSSVIITTHCS 318

RESULT 12
US-10-038-307-2
/ Sequence 2, Application US/10038307
/ Publication No. US20030134786A1
/ GENERAL INFORMATION:
/ APPLICANT: James B. ROTHMAN
/ APPLICANT: Theresa L. O'KEEFE
/ APPLICANT: Engin OZKAYNAK
/ APPLICANT: Judith J. HEALRY
/ TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
/ FILE REFERENCE: 7853-253-999
```

/ CURRENT APPLICATION NUMBER: US/10/038.307
/ CURRENT FILING DATE: 2002-06-28
/ NUMBER OF SEQ ID NOS: 26
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 333
/ TYPE: PRF
/ ORGANISM: Homo sapiens
US-10-038-307-2

Query Match 99.0%; Score 1520; DB 14; Length 333;
Best Local Similarity 100.0%; Pred. No. 2.5e-147;
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGRRRDEGPGACGCGFDLYFIIDKSGSVLHNNIYFVEQLAHKFTISPOLRMSFIVFS 60
DB 27 GGGRRRDEGPGACGCGFDLYFIIDKSGSVLHNNIYFVEQLAHKFTISPOLRMSFIVFS 86
QY 61 TRGTTLMKLTEDRQIRQGLELQKVLPGDPTVMHEGPERASQIYYENRQGYRTASVII 120
DB 87 TRGTTLMKLTEDRQIRQGLELQKVLPGDPTVMHEGPERASQIYYENRQGYRTASVII 146
QY 121 ALTDGELHEDLFFYSEREARNSRDGAIVYCVGVKDFNETQLARIADSKDHVPVNDGFQ 180
DB 147 ALTDGELHEDLFFYSEREARNSRDGAIVYCVGVKDFNETQLARIADSKDHVPVNDGFQ 206
QY 181 ALGGIHSILKSCIEILAAEPSTICAGESFQVYVVRNGGFRHARNVDRVLCSPKINDSVT 240
DB 207 ALGGIHSILKSCIEILAAEPSTICAGESFQVYVVRNGGFRHARNVDRVLCSPKINDSVT 266
QY 241 LNEKPSVEDTYLLCPAPILKEVGKAAALQVSMNDGLSFSSVITTTTHCS 292
DB 267 LNEKPSVEDTYLLCPAPILKEVGKAAALQVSMNDGLSFSSVITTTTHCS 318
RESULT 13
US-10-201-292-2
/ Sequence 2, Application US/10201292
/ Publication No. US20030144193A1
/ GENERAL INFORMATION:
/ APPLICANT: James B. ROTTMAN
/ APPLICANT: Theresa L. O'KEEFE
/ APPLICANT: Engin OZKAYNAK
/ APPLICANT: Judith J. HEALEY
/ TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
/ FILE REFERENCE: 7853-253-999
/ CURRENT APPLICATION NUMBER: US/10/201.292
/ CURRENT FILING DATE: 2003-02-14
/ NUMBER OF SEQ ID NOS: 36
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 333
/ TYPE: PRF
/ ORGANISM: Homo sapiens
US-10-201-292-2

Query Match 99.0%; Score 1520; DB 14; Length 333;
Best Local Similarity 100.0%; Pred. No. 2.5e-147;
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGRRRDEGPGACGCGFDLYFIIDKSGSVLHNNIYFVEQLAHKFTISPOLRMSFIVFS 60
DB 27 GGGRRRDEGPGACGCGFDLYFIIDKSGSVLHNNIYFVEQLAHKFTISPOLRMSFIVFS 86
QY 61 TRGTTLMKLTEDRQIRQGLELQKVLPGDPTVMHEGPERASQIYYENRQGYRTASVII 120
DB 87 TRGTTLMKLTEDRQIRQGLELQKVLPGDPTVMHEGPERASQIYYENRQGYRTASVII 146
QY 121 ALTDGELHEDLFFYSEREARNSRDGAIVYCVGVKDFNETQLARIADSKDHVPVNDGFQ 180
DB 147 ALTDGELHEDLFFYSEREARNSRDGAIVYCVGVKDFNETQLARIADSKDHVPVNDGFQ 206
QY 181 ALGGIHSILKSCIEILAAEPSTICAGESFQVYVVRNGGFRHARNVDRVLCSPKINDSVT 240

|||||
DB 207 ALGGIHSILKSCIEILAAEPSTICAGESFQVYVVRNGGFRHARNVDRVLCSPKINDSVT 266
QY 241 LNEKPSVEDTYLLCPAPILKEVGKAAALQVSMNDGLSFSSVITTTTHCS 292
DB 267 LNEKPSVEDTYLLCPAPILKEVGKAAALQVSMNDGLSFSSVITTTTHCS 318

RESULT 14
US-10-038-307-24

/ Sequence 24, Application US/10038307
/ Publication No. US20030134786A1
/ GENERAL INFORMATION:
/ APPLICANT: James B. ROTTMAN
/ APPLICANT: Theresa L. O'KEEFE
/ APPLICANT: Engin OZKAYNAK
/ APPLICANT: Judith J. HEALEY
/ TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
/ FILE REFERENCE: 7853-253-999
/ CURRENT APPLICATION NUMBER: US/10/038.307
/ CURRENT FILING DATE: 2002-06-28
/ NUMBER OF SEQ ID NOS: 26
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 24
/ LENGTH: 345
/ TYPE: PRF
/ ORGANISM: Homo sapiens
US-10-038-307-24

Query Match 99.0%; Score 1520; DB 14; Length 345;
Best Local Similarity 100.0%; Pred. No. 2.6e-147;
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGRRRDEGPGACGCGFDLYFIIDKSGSVLHNNIYFVEQLAHKFTISPOLRMSFIVFS 60
DB 27 GGGRRRDEGPGACGCGFDLYFIIDKSGSVLHNNIYFVEQLAHKFTISPOLRMSFIVFS 86
QY 61 TRGTTLMKLTEDRQIRQGLELQKVLPGDPTVMHEGPERASQIYYENRQGYRTASVII 120
DB 87 TRGTTLMKLTEDRQIRQGLELQKVLPGDPTVMHEGPERASQIYYENRQGYRTASVII 146
QY 121 ALTDGELHEDLFFYSEREARNSRDGAIVYCVGVKDFNETQLARIADSKDHVPVNDGFQ 180
DB 147 ALTDGELHEDLFFYSEREARNSRDGAIVYCVGVKDFNETQLARIADSKDHVPVNDGFQ 206
QY 181 ALGGIHSILKSCIEILAAEPSTICAGESFQVYVVRNGGFRHARNVDRVLCSPKINDSVT 240
DB 207 ALGGIHSILKSCIEILAAEPSTICAGESFQVYVVRNGGFRHARNVDRVLCSPKINDSVT 266
QY 241 LNEKPSVEDTYLLCPAPILKEVGKAAALQVSMNDGLSFSSVITTTTHCS 292
DB 267 LNEKPSVEDTYLLCPAPILKEVGKAAALQVSMNDGLSFSSVITTTTHCS 318

RESULT 15
US-10-201-292-24

/ Sequence 24, Application US/10201292
/ Publication No. US20030144193A1
/ GENERAL INFORMATION:
/ APPLICANT: James B. ROTTMAN
/ APPLICANT: Theresa L. O'KEEFE
/ APPLICANT: Engin OZKAYNAK
/ APPLICANT: Judith J. HEALEY
/ TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
/ FILE REFERENCE: 7853-253-999
/ CURRENT APPLICATION NUMBER: US/10/201.292
/ CURRENT FILING DATE: 2003-02-14
/ NUMBER OF SEQ ID NOS: 36
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 24
/ LENGTH: 345
/ TYPE: PRF
/ ORGANISM: Homo sapiens

US-10-201-292-24

Query Match 99.0%; Score 1520; DB 14; Length 345;
Best Local Similarity 100.0%; Pred. No. 2.6e-147;
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 GGGRRREDGGPACYGFDLYFILDKSGSVLHHMNEIYFVEQLAHKFI SPQLRMSFI VFS 60
      |||
      27 GGGRRREDGGPACYGFDLYFILDKSGSVLHHMNEIYFVEQLAHKFI SPQLRMSFI VFS 86
QY      61 TRGTTLMKLTEDREQIRQGLEELQKLPFGDPTVMHEGFERASBOIYYENRQGYRTASVIT 120
      |||
      87 TRGTTLMKLTEDREQIRQGLEELQKLPFGDPTVMHEGFERASBOIYYENRQGYRTASVIT 146
QY      121 ALTDGELHEDLPFYSEREANSRDIGAIYVCVGVDFNETQLARIADSKDHVPVNDGFQ 180
      |||
      147 ALTDGELHEDLPFYSEREANSRDIGAIYVCVGVDFNETQLARIADSKDHVPVNDGFQ 206
QY      181 ALQGIHSHILKKSCEIILAEPSITICAGESFQVVVRGNGFRHARNVDRVLCSPKINDSVT 240
      |||
      207 ALQGIHSHILKKSCEIILAEPSITICAGESFQVVVRGNGFRHARNVDRVLCSPKINDSVT 266
QY      241 LNEKPFVSVEDTYLCPAPILKEVGKALQVSNNDGLSFISSVIITTHCS 292
      |||
      267 LNEKPFVSVEDTYLCPAPILKEVGKALQVSNNDGLSFISSVIITTHCS 318
DB
```

Search completed: June 13, 2005, 20:36:49
Job time : 96.5881 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 19:47:08 ; Search time 21.1647 Seconds
(without alignment)
1341.095 Million cell updates/sec

Title: us-09-970-076-2_COPY_27_321

Perfect score: 1536
Sequence: 1 GCGGRREDGPGACYGDFLY.....GLSPSSSVITTTTCSDGS 295

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR_79:.*
2: pir1:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	145.5	9.5	1153	1 RWHUB	cell surface glyco
2	143.5	9.3	1163	1 RWHUB	cell surface glyco
3	143	9.3	3124	2 A40020	collagen alpha 1(X
4	141	9.2	1153	2 S00551	leukocyte surface
5	139	9.0	724	2 A48569	antigen Em100 - E1
6	137	8.9	1747	2 A45974	collagen alpha 1(X
7	134.5	8.8	712	2 A45638	immunodominant mic
8	131	8.5	760	1 C2MS	classical-compleme
9	130	8.5	1857	2 S31212	collagen alpha 1(X
10	130	8.5	1888	2 S78476	collagen alpha 1(X
11	128	8.3	3051	2 S42373	hypothetical prote
12	118	7.7	929	2 I51027	type XII collagen
13	113.5	7.4	496	2 A37979	cartilage matrix p
14	113.5	7.4	764	1 B8HU	complement factor
15	111.5	7.3	2944	2 A54849	collagen alpha 1(V
16	110.5	7.2	761	1 B8MS	complement factor
17	109.5	7.1	1029	1 S21369	collagen alpha 2(V
18	108.5	7.1	500	2 S66522	cartilage matrix p
19	107.5	7.0	2	A33809	cartilage matrix p
20	105.5	6.9	567	2 T28797	hypothetical prote
21	104	6.8	747	2 I51579	complement factor
22	100.5	6.5	574	2 A46283	sporozoite surface
23	100	6.5	460	2 T23087	hypothetical prote
24	99.5	6.5	597	2 S33578	rop protein - fru1
25	99	6.4	559	2 S04531	thrombospondin-rel
26	99	6.4	896958	2	dnak protein (heat
27	99	6.4	817	2 T50240	kinesin-like prote
28	99	6.4	917	2 S09646	collagen alpha 2(V
29	99	6.4	1018	1 CGH02A	collagen alpha 2(V

30	98.5	6.4	537	2 T04822	hypothetical prote
31	98	6.4	741	2 T46488	hypothetical prote
32	98	6.4	3176	2 CGH03A	collagen alpha 3(V
33	97.5	6.3	642	2 H81185	dnak protein NMB05
34	97	6.3	3137	2 A37797	collagen alpha 3(V
35	96.5	6.3	676	2 T47637	hypothetical prote
36	96	6.2	272	2 A55348	hypothetical prote
37	96	6.2	689	2 F84811	integrin alpha-1 -
38	95.5	6.2	334	2 AF1166	probable retroelem
39	95.5	6.2	932	2 JC5953	transcription regu
40	95.5	6.2	1179	2 A53213	inter-alpha-inhibi
41	94.5	6.2	1151	2 A53213	integrin alpha-B c
42	94	6.1	843	2 A40970	undulin 1 - human
43	93	6.1	642	2 E70121	hypothetical prote
44	92.5	6.0	642	2 B81917	probable chaperone
45	92	6.0	292	2 B83736	transposase (04) B

ALIGNMENTS

RESULT 1

RWHUB

cell surface glycoprotein CD11b precursor [validated] - human
N/Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein Mac
eukocyte integrin alpha chain; neutrophil adherence receptor alphas chain
C/Species: Homo sapiens (man)

C/Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 09-Jul-2004
C/Accession: A31108; A28915; A41600; A30892; A32218; A46526; A26091; I52567
R/Corbi, A.L.; Kishimoto, T.K.; Miller, U.U.; Springer, T.A.
J. Biol. Chem. 265, 12403-12411, 1988

U. Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11
B.

A/Reference number: A31108; MUID:88315033; PMID:2457584

A/Accession: A31108

A/Molecule type: mRNA

A/Residues: 1-1153 <COR>

A/Cross-references: UNIPROT:P11215; GB:J03925; NID:G187284; PIDN:AAA59544.1; PID:G307148

A/Note: part of this sequence was confirmed by protein sequencing

R/Arnaut, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.
J. Cell Biol. 106, 2153-2158, 1988

U. Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor Mac

A/Reference number: A28915; MUID:88257215; PMID:2454931

A/Accession: A28915

A/Molecule type: mRNA

A/Residues: 1-499, 501-965, 'P', 967-1153 <ARN>

A/Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:G186935; PIDN:AAA594;

A/Note: the authors translated the codon TAC for residue 1129 as Thr

A/Note: part of this sequence, including the amino end of the mature protein, was confirm

R/Shelley, C.S.; Arnaut, M.A.
Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991

U. Title: The promoter of the CD11b gene directs myeloid-specific and developmentally regu

A/Reference number: A41600; MUID:92073318; PMID:1683702

A/Accession: A41600

A/Molecule type: DNA

A/Residues: 1-9 <SHB>

A/Cross-references: GB:M76724; NID:G180018; PIDN:AAA54410.1; PID:G553215

R/Arnaut, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.
Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988

U. Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhesic

A/Reference number: A94193; MUID:88190151; PMID:2833753

A/Accession: A30892

A/Molecule type: mRNA

A/Residues: 917-1042 <AR2>

A/Cross-references: GB:M18044

R/Hickstein, D.D.; Hickey, M.U.; Ozols, J.; Baker, D.M.; Beck, A.L.; Roth, G.J.
Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989

U. Title: cDNA sequence for the alpha subunit of the human neutrophil adherence receptor

A/Reference number: A32218; MUID:89088893; PMID:2563162

A/Accession: A32218

A/Molecule type: mRNA

A/Residues: 9-1153 <HIC>

A/Cross-references: GB:J04145; NID:G189068; PIDN:AAA59903.1; PID:G386975

A>Note: part of this sequence was confirmed by protein sequencing
R:Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.
J.Immunol. 150, 480-490, 1993

A>Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-III_b during evolution.

A:Reference number: A46526; MUID:93123748; PMID:8419480

A:Accession: A46526

A>Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-499,501-1153 <PE>

A:Cross-references: GB:S52227; NID:g263047; PIDN:AAB24821.1; PID:g563049

A>Note: the last three bases of intron 13, CAG, are included in some but not all mature R.NA sequences extracted from NCBI backbone (NCBIR:121963)

R:Pierce, M.W.: Remold-O'Donnell, E.; Todd III, R.F.; Arnaut, M.A.
Biochim. Biophys. Acta 874, 368-371, 1986

A>Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across species

A:Reference number: A90664; MUID:87076671; PMID:3539202

A:Accession: A26091

A:Molecule type: protein

A:Residues: 17-31<PI>

A:Experimental source: granulocytes

R:Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G.
Blood 79, 865-870, 1992

A>Title: Characterization of the myeloid-specific CD11b promoter.

A:Reference number: 152567; MUID:92144986; PMID:1346576

A:Accession: 152567

A>Status: translated from GB/EMBL/DDBJ

A:Molecule type: RNA

A:Residues: 1-9 <RES>

A:Cross-references: GB:M64477; NID:g160184; PIDN:AAAS1960.1; PID:g553219

C.Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1

C:Genetics:

A:Gene: GDB:iTGM; CR3A

A:Cross-references: GDB:120599; OMIM:120980

A:Map position: 16p11.2-16p11.2

A>Note: promotor contains a GATA motif and two Sp1 consensus binding sites

C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homolog

C:Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; mag

F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-1108/Product: cell surface glycoprotein CD11b #status experimental <MNT>
F:148-318/Domain: extracellular #status predicted <EXT>
F:465-473/Region: calcitum/magnesium binding #status predicted <VMA>
F:530-538/Region: calcitum/magnesium binding #status predicted
F:593-601/Region: calcitum/magnesium binding #status predicted
F:1109-1134/Domain: transmembrane #status predicted <TMW>
F:1135-1153/Domain: intracellular #status predicted <INT>
P:86,240,391,469,693,697,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Binding

N:Alternate names: leukocyte adhesion receptor p150,95 alpha chain
C:Species: Homo sapiens (man)
C>Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text_change 09-Jul-2004
C:Accession: A36584; A35543; S00864
R:Cordt, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 12750-12751, 1990
A:Reference number: A36584
A:Contents: erratum
A:Accession: A36584
A:Molecule type: DNA
A:Residues: 1-1163 <COR>
A:Cross-references: UNIPROT:P20702
A>Note: This revision to the sequence from reference A35543 includes the carboxyl end
R:Cordt, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 2782-2788, 1990
A>Title: Genomic structure of an integrin alpha subunit, the leukocyte p150,95 molecule,
A:Reference number: A35543; MUID:90153906; PMID:2303426
A:Accession: A35543
A:Molecule type: DNA
A:Residues: 1-834 <CO2>
A>Note: This sequence has been revised in reference A36584
R:Cordt, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.
EMBO J. 6, 4023-4028, 1987
A>Title: CDNA cloning and complete primary structure of the alpha subunit of a leukocyte
A:Reference number: S00864; MUID:88166645; PMID:3327687
A:Accession: S00864
A:Molecule type: mRNA
A:Residues: 1-755,'L',757-1163 <CO3>
A:Cross-references: GB:M81699; EMBL:Y00093; NID:G487829; PIDN:AAA59180.1; PID:G487830
A:Note: part of this sequence was confirmed by protein sequencing
C:Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on mye
C:Genetics:
A:Gene: GDB:ITGAX; CD11C
A:Cross-references: GDB:119758; OMIM:151510
A:Map position: 16p11.2-16p11.2
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homol
C:Keywords: calcium; cell adhesion; glycoprotein; heterodimer; magnesium; tandem repeat;
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-1163/Product: cell surface glycoprotein CD11c #status predicted <MAT>
F:20-1107/Domain: extracellular #status predicted <EXT>
F:149-319/Domain: von Willebrand factor type A repeat homology <VMA4>
F:1108-1133/Domain: transmembrane #status predicted <TM>
F:1134-1163/Domain: intracellular #status predicted <INT>
F:61,89,392,697,735,899,939,1050/Binding site: carbohydrate (Asn) (covalent) #status pre

Query Match 9.3%; Score 143.5; DB 1; Length 1163;
Best Local Similarity 24.8%; Pred. No. 0.0024;
Matches 55; Conservative 43; Mismatches 89; Indels 35; Gaps 10;

Dy 18 DLYFLDKSGSV-LHHMNEIYFPEDLAHKFIISPOLRMSEFYVSTGTTLMLKLTEDRE-- 74
|::|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 151 DIVFLPIDSGLSSISHNFALPMNVFAVVISOFORPSFOFSLMOPSNKFOTFTPEEFRTS 210
75 ---QIRGLEELIKYVLPGSDTYMHGEFERASERQIYYENNGVYTAS-VIALTDGSLHED 130
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Dy 211 NPUSLIASHHQIQ-----GFYTTATRIQNVAHLPLFASLGARDATKILLVITDGKKEGD 265
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 131 LEFGSEREARNRDIGAIYVCVG-----KDFNETOLARIAD--SKDHVPFVNDG 178
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
266 SLDDYKD-VIAPMDAGIIIFALGVGLAFQNRMSWEKLD-----ISKPEQHIFKYED- 318
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Dy 179 FOALQGIIHSIIKKSCIEILLAEPTTCAGESFOVVVRGNPF 220
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 319 FDLAKDIQWLKER-----IFALEGFTTSSSFLEMAQEGF 356
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 3
A40020 collagen alpha 1(XII) chain precursor - chicken
N:Alternate names: fibrochimerin
C:Species: Gallus gallus (chicken)
C>Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: A40020; A34485; B34485; A28037; S23814; S22524; S28811

R.Yamagata, M.; Yamada, K.M.; Yamada, S.S.; Shimomura, T.; Tanaka, H.; Nishida, Y.; Obata, U. Cell Biol. 115, 209-221, 1991
 A>Title: The complete primary structure of type XII collagen shows a chimeric molecule with a novel region of type IX collagen, and short collagenous domains with an Arg-Gly-Asp site.
 A/Reference number: A40020; MUID:92011862; PMID:1918137
 A/Accession: A40020
 A/Molecule type: mRNA
 A/Residues: 1-3124 <YAM>
 A/Cross-references: UNIPROT:P13944; GB:D00824; NID:9222810; PIDN:BAA00701.1; PID:9222811
 A/Note: In the authors' translation residues 1216-1219 are shown after residue 1235 and, R.Gordon, M.K.; Garecke, D.R.; Dublet, B.; van der Rest, M.; Olsen, B.R.
 J. Biol. Chem. 264, 19772-19778, 1989
 A>Title: Type XII collagen. A large multidomain molecule with partial homology to type I
 A/Reference number: A34485; MUID:90062079; PMID:2584192
 A/Accession: A34485
 A/Molecule type: mRNA
 A/Residues: 2456-2756, 'A', 2760-2802, 'F', 2804-2976, 'F', 2978-3124 <GOR>
 A/Cross-references: EMBL:U05137; NID:9211284; PIDN:AAA8635.1; PID:9211285
 A/Accession: B34485
 A/Molecule type: protein
 A/Residues: 2772-2792;2846-2873 <GOR2>
 R.Gordon, M.K.; Garecke, D.R.; Olsen, B.R.
 Proc. Natl. Acad. Sci. U.S.A. 86, 6040-6044, 1987
 A>Title: Type XII collagen: distinct extracellular matrix component discovered by cDNA
 A/Reference number: A28037; MUID:87317590; PMID:3476925
 A/Accession: A28037
 A/Molecule type: mRNA
 A/Residues: 2960-2976, 'F', 2978-3074, 'AG' <GOR3>
 A/Cross-references: EMBL:M17375; NID:9211649; PIDN:AAA8718.1; PID:9211650
 A/Note: This sequence has been revised in reference A34485
 R.Koch, M.; Bernasconi, C.; Chiquet, M.
 Eur. J. Biochem. 207, 847-856, 1992
 A>Title: A major oligomeric fibroblast proteoglycan identified as a novel large form of
 A/Reference number: 523814; MUID:92362621; PMID:1332460
 A/Accession: 523814
 A/Molecule type: protein
 A/Residues: 'X', 1333, 'Q', 1335-1347;1914-1928;2504, 'X', 2506, 'X', 2508-2511, 'X', 2513-2517 <R>
 J. Dublet, B.; van der Rest, M.
 J. Biol. Chem. 267, 17724-17727, 1992
 A>Title: Type XII collagen is expressed in embryonic chick tendons. Isolation of peptide
 A/Reference number: 522254; MUID:88087065; PMID:3121603
 A/Accession: 522254
 A/Molecule type: protein
 A/Residues: 2831-2832, 'T', 2834, 'R', 2836-2843;3002-3014 <DUB>
 R.Tueb, J.; Tueb, B.
 Biochim. Biophys. Acta 1171, 97-98, 1992
 A>Title: The two splice variants of collagen XII share a common 5' end.
 A/Reference number: 528811; MUID:93042014; PMID:1420368
 A/Accession: 528811
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-24, 1189-1257, 'G', 1259-1263, 'E', 1265-1280 <TRU>
 A/Cross-references: EMBL:X67327
 C/Genetics:
 A/Intons: 2845/3; 2863/3; 2887/3; 2899/3; 2922/1; 2985/1; 3008/1; 3065/1
 C/Keywords: alternative splicing; cell binding; coll binding; coll; connective tissue; disulfide
 P:1-23/Domin: signal sequence #status predicted <SIG>
 P:24-3124/Product: collagen alpha 1(XII) chain #status predicted <MAT>
 F:24,1189-3124/Product: collagen alpha 1(XII) chain short splice form #status predicted
 F:24-114/Domin: IIA #status predicted <IIA>
 F:24-105/Domin: fibronectin type III repeat homology <FN3A>
 F:137-301/Domin: von Willebrand factor type A repeat homology <VMA1>
 F:1332-425/Domin: IIB #status predicted <IIB>
 F:1332-414/Domin: fibronectin type III repeat homology <FN3B>
 F:437-601/Domin: von Willebrand factor type A repeat homology <VMA2>
 F:629-1178/Domin: IIC #status predicted <IIC>
 F:630-711/Domin: fibronectin type III repeat homology <FN3C>
 F:721-802/Domin: fibronectin type III repeat homology <FN3D>
 F:812-895/Domin: fibronectin type III repeat homology <FN3E>
 F:905-986/Domin: fibronectin type III repeat homology <FN3F>
 F:995-1076/Domin: fibronectin type III repeat homology <FN3G>
 F:1086-1169/Domin: fibronectin type III repeat homology <FN3H>
 F:1197-1361/Domin: von Willebrand factor type A repeat homology <VMA3>

F:1384-2295/Domin: IID #status predicted <IID>
 F:1384-1465/Domin: fibronectin type III repeat homology <FN3I>
 F:1474-1557/Domin: fibronectin type III repeat homology <FN3J>
 F:1566-1647/Domin: fibronectin type III repeat homology <FN3K>
 F:1555-1738/Domin: fibronectin type III repeat homology <FN3L>
 F:1756-1838/Domin: fibronectin type III repeat homology <FN3M>
 F:1847-1928/Domin: fibronectin type III repeat homology <FN3N>
 F:1937-2019/Domin: fibronectin type III repeat homology <FN3O>
 F:2028-2110/Domin: fibronectin type III repeat homology <FN3P>
 F:2119-2199/Domin: fibronectin type III repeat homology <FN3Q>
 F:2225-2294/Domin: fibronectin type III repeat homology <FN3R>
 F:2225-2490/Domin: von Willebrand factor type A repeat homology <VMA4>
 F:2438-2440/Domin: cell adhesion #status predicted
 F:2509-2750/Domin: IXP, homologous to NC4 domain of type IX collagen #status predicted
 F:2751-2902/Domin: collagenous COL2 #status predicted <COL2>
 F:2899-2901/Domin: cell attachment (R-G-D) motif
 F:2903-2945/Domin: non-collagenous NC2 #status predicted <NC2>
 F:2946-3048/Domin: collagenous COL1 #status predicted <COL1>
 F:3049-3124/Domin: non-collagenous NC1 #status predicted <NC1>
 F:32,1006,1032,1044,1512,1767,2210,2273,2532,2683/Binding site: carbohydrate (Asn) (cova)
 F:2780,2789,2836,2842,2860,2866,2869,3004,3007/Modified site: hydroxyproline (Pro) #status
 Query Match 9.3%; Score 143; DB 2; Length 3124;
 Best Local Similarity 26.1%; Pred. No. 0.0098;
 Matches 65; Conservative 51; Mismatches 97; Indels 36; Gaps 14;
 QY 18 DLYFLDKSGSV-LHHNNEIYFVEOLAHKF-ISP-OLRMSFIYSTRGTLMLKLTEDRE 74
 Db 439 DVVFPLVGSISIGIANVKKRAFLFVLVKSFEISPRKQVSLVOYSR--DPRHESLNRY 496
 QY 75 QIROGLEBELQKLP--GGDTYMEGFERASEQIYENRQGR--TASVIALTDGELHED 130
 Db 497 NRVYDIIQAIINTFPYRGSGTNTGKAMTYAEKVFVTSK--GSRPVPRVMTLITDGK--SSD 554
 QY 131 LFFYSREARSRDGLAIVYCVGKDPENFQRLIAD--SKDRVFPNDGFQALQGIHS 188
 Db 555 AF--KSPAKILKADVAEIPAVGKQAVRLELAISPPAEHTVYTED--FDLQRIISFE 610
 QY 189 ILKSCIEIIAEPSTICAGESFQVVRGNGFRRARV--DRYLCSFKINDSVTLNEKP 245
 Db 611 LTQSVLCIRT-----EQELAIKRSYPAKAMVSDVTSDSFKWSMAGSE-- 657
 QY 246 FSVEDTYLL 254
 Db 658 ---EKSYLI 663
 RESULT 4
 500551
 leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse
 N:Alternate names: complement-3 receptor alpha chain
 C:Species: Mus musculus (house mouse)
 C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
 C/Accession: S00551; I59078
 R:Pyteia, R.
 EMBO J. 7, 1371-1378, 1988
 A>Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the
 A/Reference number: 500551; MUID:88312584; PMID:3044779
 A/Accession: S00551
 A/Molecule type: DNA
 A/Residues: 1-1153 <PYT>
 A/Cross-references: UNIPROT:P05555; EMBL:X07640; NID:952982; PIDN:CAA30479.1; PID:952983
 A/Note: the authors translated the codon CAC for residue 569 as Gln
 R:Satte, L.; Roman, J.M.; Teplov, D.B.; Dreyer, W.J.; Gee, C.E.; Larson, R.S.; Roberts,
 Proc. Natl. Acad. Sci. U.S.A. 83, 5644-5648, 1986
 A>Title: A partial genomic DNA clone for the alpha subunit of the mouse complement recep
 A/Reference number: I59078; MUID:86287312; PMID:2942940
 A/Accession: I59078
 A/Status: preliminary; translated from GB/EMBL/DDb
 A/Molecule type: DNA
 A/Residues: 11-44 <RES>
 A/Cross-references: GB:M14293; NID:9198993; PIDN:AAA39484.1; PID:9554193
 C/Genetics:

A/Gene: Mac-1
C/Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
C/Keywords: cell adhesion; glycoprotein; transmembrane protein
F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-1153/Product: leukocyte surface glycoprotein Mac-1 alpha chain #status experimental
F:148-318/Domain: von Willebrand factor type A repeat homology <VWA2>
F:1106-1129/Domain: transmembrane #status predicted <TM>

Query Match 9.2%; Score 141; DB 2; Length 1153;
Best Local Similarity 24.2%; Pred. No. 0.0038;
Matches 71; Conservative 50; Mismatches 96; Indels 76; Gaps 16;

Qy 18 DLVFIIDKSGSVLHHMNEIYFVEQLAHKFIS---POLRMSFVSTRGTTLMKLTED-- 72
Db 150 DIVFLIDSGSGSI---NNIDF---QKMEFVSTWEGKSKSTLFS-----LMQYSDFR 197
Qy 73 -----REGIROGLBELQKLP-----GGDTYMHGEGERASEQIYYE--NNGRTASVIALT 123
Db 198 IHFTFNDPKRNPSPRSHSPKOLNGRTKTASGIRKVVRELPHKTNGARENAKILVIT 257
Qy 124 DGELEHEDLPFYSE--REANRGRDGAIVYCVGVND-FNETQIARIAD-----SKDHVP 174
Db 258 DGEKFGDPLDYKDVIPEDRA---GVIRYVIGVGNAPKPKPSRRELDTISKAPGEHVPQ 314
Qy 175 VNDGFQALQGIHSLKSKCIEILAEPSTICAGESFQVVGNGGFRHARNVDRVLSFK 234
Db 315 V-DNFEALNTIONLOEK---IPALIEGTGTSTSFHEHMSGEGF----- 355
Qy 235 INDVTLNEKPSVEDTYLCPAFLKEVG---MKALQVSNMDSGISISS 283
Db 356 -SASITSN-----GPILSGVSGFDMAGAFLYTSKDKVFIINTT 393

RESULT 5

A48569
antiGen Em100 - Elmeria maxima
C/Species: Elmeria maxima
C/Date: 01-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C/Accession: A48569
R:Pasamontes, L.; Hug, D.; Humbelin, M.; Weber, G.
Mol. Biochem. Parasitol. 57, 171-174, 1993
A/Title: Sequence of a major Elmeria maxima antigen homologous to the Elmeria tenella m
A/Reference number: A48569; MUID:93149203; PMID:842611
A/Accession: A48569
A/Status: preliminary
A/Molecule type: nucleic acid
A/Residues: 1-724 <PAS>
A/Cross-references: UNIPROT:Q04588; GB:M99058; NID:9158890; PID:9158891
A/Note: sequence extracted from NCBI backbone (NCBIN:123776, NCBI:P:123777)
F:45-218/Domain: von Willebrand factor type A repeat homology <VWA2>
F:238-296/Domain: thrombospondin type 1 repeat homology <THR1>
F:309-371/Domain: thrombospondin type 1 repeat homology <THR2>
F:372-432/Domain: thrombospondin type 1 repeat homology <THR3>
F:433-493/Domain: thrombospondin type 1 repeat homology <THR4>
F:494-556/Domain: thrombospondin type 1 repeat homology <THR5>
F:560-610/Domain: thrombospondin type 1 repeat homology <THR6>

Query Match 9.0%; Score 139; DB 2; Length 724;
Best Local Similarity 26.6%; Pred. No. 0.003; Indels 34; Gaps 14;
Matches 59; Conservative 40; Mismatches 89;

Qy 13 CYGGFDLVFIIDKSGSV-LHHMNEIYFVEQLAHKF-ISP-QLRMSFVSTRGTTLMKL 69
Db 42 CTRLLDVVLVYDESGISTSYKVRSFISNFAQMPLSPDVAVGVLTFGTSVATMDL 101
Qy 70 TEDREQIRQGLEBLQKLP--GGDTYMHGEGERASEQIYYENRQCYR--TASVIALTDG 125
Db 102 SDSRAQNDLLAAAKKLPAAGSTYTHGLAKA-EEILFFQKGGGRNAPKMLVWTDG 160
Qy 126 ELHEDLPFYSE-----ANRSDIGAIVYCVGV-KDFNETOLARIA--DSKHV-PP- 174
Db 161 A-----SSRSQTLAAEKLNRGVIIYVGLGVGVSNAECRSINGCDTSTVECPR 212

Qy 175 -VNDGFQALQGIHSLKSKCIEI---LAAEPSTI--CAGE 209
Db 213 YLQSNMGVSSQINGIILKAKDLADAVCSWESEYGPCEG 254

RESULT 6
A45974
collagen alpha 1(XIV) chain precursor, short form 2 - chicken

N/Alternate names: undulin
C/Species: Gallus gallus (chicken)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C/Accession: A45974; S30085; S22916; S17035; S20833
R:Gercke, D.R.; Foley, J.W.; Caetagnola, P.; Gennari, M.; Dublet, B.; Canceda, R.; Lin
J. Biol. Chem. 268, 12177-12184, 1993
A/Title: Type XIV collagen is encoded by alternative transcripts with distinct 5' regions
ns.

A/Reference number: A45974; MUID:93280195; PMID:8505337
A/Accession: A45974
A/Status: preliminary
A/Molecule type: mRNA; protein
A/Residues: 1-1747 <GER>
A/Cross-references: UNIPROT:P32018
A/Experimental source: embryo skin
A/Note: sequence inconsistent with the nucleotide translation
A/Note: sequence extracted from NCBI backbone (NCBIN:133364, NCBI:P:133365)
R:Apce, S.S.
submitted to the EMBL Data Library, March 1992
A/Reference number: S30085
A/Accession: S30085

A/Molecule type: mRNA
A/Residues: 1472-1660 <APT>
A/Cross-references: EMBL:X65122; NID:962871; PID:CAA46238.1; PID:9938175
R:Trueb, J.; Trueb, B.
Eur. J. Biochem. 207, 549-557, 1992
A/Title: Type XIV collagen is a variant of undulin.

A/Reference number: S22916; MUID:92359443; PMID:1339349
A/Accession: S22916
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 286-494, 'Q', 496-834, 'A', 836-1119, 'KL', 1122-1402, 1409-1439 <TRU>
R:Gordon, M.K.; Caetagnola, P.; Dublet, B.; Linsemayer, T.F.; van der Rest, M.; Mayne, F
Eur. J. Biochem. 201, 333-338, 1991
A/Title: Cloning of a cDNA for a new member of the class of fibril-associated collagens v
A/Reference number: S17035; MUID:92037585; PMID:1935930
A/Accession: S17035

A/Molecule type: mRNA
A/Residues: 1472-1659 <GOR1>
A/Accession: S20833
A/Molecule type: protein
A/Residues: 1551-1570; 1593-1599; 1639-1667 <GOR2>
C/Keywords: alternative splicing; coiled coil; extracellular matrix; glycoprotein; trimer

F:40-204/Domain: von Willebrand factor type A repeat homology <VWA1>
F:236-317/Domain: fibronectin type III repeat homology <FN3>
F:326-409/Domain: fibronectin type III repeat homology <FN3B>
F:418-498/Domain: fibronectin type III repeat homology <FN3C>
F:507-591/Domain: fibronectin type III repeat homology <FN3D>
F:625-707/Domain: fibronectin type III repeat homology <FN3E>
F:716-798/Domain: fibronectin type III repeat homology <FN3F>
F:806-893/Domain: fibronectin type III repeat homology <FN3G>
F:924-1089/Domain: von Willebrand factor type A repeat homology <VWA2>
F:1111-1352/Domain: non-collagenous NC4 #status predicted <NC4>
F:1511-1553/Domain: non-collagenous NC2 #status predicted <NC2>
F:1554-1659/Domain: triple helical domain COL1 #status predicted <COL1>

Query Match 8.9%; Score 137; DB 2; Length 1747;
Best Local Similarity 25.4%; Pred. No. 0.014; Indels 26; Gaps 12;
Matches 63; Conservative 45; Mismatches 114;

Qy 18 DLVFIIDKSGSV-LHHMNEIYFVEQL--AHKISFQ-LRMSFVSTRGTTLMKL--E 71
Db 926 DLVFIIDKSGSVSDDDPNKISFLYSTVGALMDIGDGVVALIQSDDDRTFKLNAYK 985
Qy 72 DREQIRQGLEBLQKLPVGGDTYMHGEGERASEQIYYENRQCYR--TASVIALTDGELHE 129

Db 966 TKTLLLEIQQI--AYKGNNTKGAIAIKHARE-VLFTSEAGRRKCIPIKVLVITIDGRSOD 1042
Oy 130 DLFFYSSEARNRSL-GAIYVCVGDENETOLARIAD--SKDHVPVNDGFOALQGI 186
Db 1043 DV-----NKVRREMLDGFSPPAIVADADYSELVNTSKSERVAVDD-FDAFTKIE 1096
Oy 187 HSILKSCIRILAAEPSTICAGESFQVAVVNGGFRHAAVNDVLCSPKINDSVTLNEKPF 246
Db 1097 DELITFVCETASATCLPLVKQDNFA-----GFGMGEMFGLVEKFEALDGVSMPEPTF 1150
Oy 247 SVEDTYLL 254
Db 1151 NVYPCYRL 1158

RESULT 7

A45638
Immunodominant microneme protein Exp100 - Eimeria tenella
C/Species: Eimeria tenella
C/Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C/Accession: A45638
R/Tonley, F.M.; Clarke, L.E.; Kawazoe, U.; Djikema, R.; Kok, J.J.
Mol. Biochem. Parasitol. 49, 277-288, 1991
A/Title: Sequence of the gene encoding an immunodominant microneme protein of Eimeria tenella
A/Reference number: A45638; MUID:92131064; PMID:1775171
A/Accession: A45638
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-712 <TOM>
A/Cross-references: UNIPROT:O43981; GB:AFO32905; GB:M73495; NID:G2707732; PIDN:AAD03350.
A/Note: Sequence extracted from NCBI backbone (NCBI:77752, NCBI:77756)
F/48-218/Domain: von Willebrand factor type A repeat homology <VWA1>
F/238-296/Domain: thrombospondin type 1 repeat homology <THR1>
F/309-371/Domain: thrombospondin type 1 repeat homology <THR2>
F/372-432/Domain: thrombospondin type 1 repeat homology <THR3>
F/433-493/Domain: thrombospondin type 1 repeat homology <THR4>
F/494-556/Domain: thrombospondin type 1 repeat homology <THR5>
F/560-610/Domain: thrombospondin type 1 repeat homology <THR6>

Query Match 8.8%; Score 134.5; DB 2; Length 712;
Best Local Similarity 24.6%; Pred. No. 0.0068;
Matches 55; Conservative 40; Mismatches 88; Indels 41; Gaps 11;
Oy 13 CYGGFDLYFLIDKSGSV-LHHMNEIYFVEQLAHKF-ISPO-LRMSFVSTRTGTLMLKL 69
Db 45 CTSLIDWLVLVDSSGISTSNFRKVRQPIEDPVNSMPISPEVARKGLITFT----- 96
Oy 70 TEDREQIRQGLEBLQKVLPG-----GDTVMHGFERRASEQIYVENRQYR-- 114
Db 97 ---RSKVRWMLSDPKATVPSLAISAARLSYSTGVTTYHGLQDA-KKLVDTNAGANN 152
Oy 115 TASVITALTDELHEDLFFYSEREARNSRDIGAIYVCV-KDNETOLARIADSKDHV 173
Db 153 VPKVLVLVWTDAA--SNLPSQTRSSAAALRDAGAIYVVLGVSGVNSSECRSIAGCSTNC 210
Oy 174 P-----VNDGFQALQGIHSLKSKSCIELAEPSTI--CAGE 209
Db 211 PRYLOSMKSNVNTQOVNGIILKACDLADAVCSEWSEYGPVGE 254

RESULT 8

C2MS
classical-complement-pathway C3/C5 convertase (EC 3.4.21.43) C2 component precursor - mc
N/Alternate names: C3 convertase; C5 convertase; complement C2
C/Species: Mus musculus (house mouse)
C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C/Accession: A38876; B36593; I54429
R/Ishikawa, N.; Nonaka, M.; Wetzel, R.A.; Colten, H.R.
submitted to GenBank, January 1991
A/Reference number: A38875
A/Accession: A38876
A/Molecule type: DNA

A/Residues: 1-760 <IS2>
A/Cross-references: UNIPROT:P21180; GB:M57891; GB:J05661; NID:g192436; PIDN:AAA63294.1; R/Ishikawa, N.; Nonaka, M.; Wetzel, R.A.; Colten, H.R.
J. Biol. Chem. 265, 19040-19046, 1990
A/Title: Murine complement C2 and factor B genomic and cDNA cloning reveals different met
A/Reference number: A36593; MUID:91035430; PMID:2229060
A/Accession: B36593

A/Molecule type: mRNA
A/Residues: 1-760 <ISH>
A/Cross-references: EMBL:M57891; NID:g192436; PIDN:AAA63294.1; PID:g192437
R/Falus, A.; Wakeland, E.K.; McConnell, T.J.; Gitlin, J.; Whitehead, A.S.; Colten, H.R.
Immunogenetics 25, 290-298, 1987
A/Title: DNA polymorphism of MHC III genes in inbred and wild mouse strains.
A/Reference number: I54429; MUID:87192938; PMID:2883115
A/Accession: I54429

A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 660-677, R', 679, 681-723, G', 725 <RBS>
A/Cross-references: GB:M16271; NID:g199289; PIDN:AAA39562.1; PID:g199290
C/Genetics:

A/Introns: 16/1; 91/1; 153/1; 212/1; 245/3; 290/3; 337/1; 384/1; 414/1; 461/1; 492/3; 532/3
A/Complex: The proenzyme forms a complex with C4a and is activated by cleavage into C2a, C3, C4, C5, C6, C7, C8, C9, C10, C11, C12, C13, C14, C15, C16, C17, C18, C19, C20, C21, C22, C23, C24, C25, C26, C27, C28, C29, C30, C31, C32, C33, C34, C35, C36, C37, C38, C39, C40, C41, C42, C43, C44, C45, C46, C47, C48, C49, C50, C51, C52, C53, C54, C55, C56, C57, C58, C59, C60, C61, C62, C63, C64, C65, C66, C67, C68, C69, C70, C71, C72, C73, C74, C75, C76, C77, C78, C79, C80, C81, C82, C83, C84, C85, C86, C87, C88, C89, C90, C91, C92, C93, C94, C95, C96, C97, C98, C99, C100, C101, C102, C103, C104, C105, C106, C107, C108, C109, C110, C111, C112, C113, C114, C115, C116, C117, C118, C119, C120, C121, C122, C123, C124, C125, C126, C127, C128, C129, C130, C131, C132, C133, C134, C135, C136, C137, C138, C139, C140, C141, C142, C143, C144, C145, C146, C147, C148, C149, C150, C151, C152, C153, C154, C155, C156, C157, C158, C159, C160, C161, C162, C163, C164, C165, C166, C167, C168, C169, C170, C171, C172, C173, C174, C175, C176, C177, C178, C179, C180, C181, C182, C183, C184, C185, C186, C187, C188, C189, C190, C191, C192, C193, C194, C195, C196, C197, C198, C199, C200, C201, C202, C203, C204, C205, C206, C207, C208, C209, C210, C211, C212, C213, C214, C215, C216, C217, C218, C219, C220, C221, C222, C223, C224, C225, C226, C227, C228, C229, C230, C231, C232, C233, C234, C235, C236, C237, C238, C239, C240, C241, C242, C243, C244, C245, C246, C247, C248, C249, C250, C251, C252, C253, C254, C255, C256, C257, C258, C259, C260, C261, C262, C263, C264, C265, C266, C267, C268, C269, C270, C271, C272, C273, C274, C275, C276, C277, C278, C279, C280, C281, C282, C283, C284, C285, C286, C287, C288, C289, C290, C291, C292, C293, C294, C295, C296, C297, C298, C299, C300, C301, C302, C303, C304, C305, C306, C307, C308, C309, C310, C311, C312, C313, C314, C315, C316, C317, C318, C319, C320, C321, C322, C323, C324, C325, C326, C327, C328, C329, C330, C331, C332, C333, C334, C335, C336, C337, C338, C339, C340, C341, C342, C343, C344, C345, C346, C347, C348, C349, C350, C351, C352, C353, C354, C355, C356, C357, C358, C359, C360, C361, C362, C363, C364, C365, C366, C367, C368, C369, C370, C371, C372, C373, C374, C375, C376, C377, C378, C379, C380, C381, C382, C383, C384, C385, C386, C387, C388, C389, C390, C391, C392, C393, C394, C395, C396, C397, C398, C399, C400, C401, C402, C403, C404, C405, C406, C407, C408, C409, C410, C411, C412, C413, C414, C415, C416, C417, C418, C419, C420, C421, C422, C423, C424, C425, C426, C427, C428, C429, C430, C431, C432, C433, C434, C435, C436, C437, C438, C439, C440, C441, C442, C443, C444, C445, C446, C447, C448, C449, C450, C451, C452, C453, C454, C455, C456, C457, C458, C459, C460, C461, C462, C463, C464, C465, C466, C467, C468, C469, C470, C471, C472, C473, C474, C475, C476, C477, C478, C479, C480, C481, C482, C483, C484, C485, C486, C487, C488, C489, C490, C491, C492, C493, C494, C495, C496, C497, C498, C499, C500, C501, C502, C503, C504, C505, C506, C507, C508, C509, C510, C511, C512, C513, C514, C515, C516, C517, C518, C519, C520, C521, C522, C523, C524, C525, C526, C527, C528, C529, C530, C531, C532, C533, C534, C535, C536, C537, C538, C539, C540, C541, C542, C543, C544, C545, C546, C547, C548, C549, C550, C551, C552, C553, C554, C555, C556, C557, C558, C559, C560, C561, C562, C563, C564, C565, C566, C567, C568, C569, C570, C571, C572, C573, C574, C575, C576, C577, C578, C579, C580, C581, C582, C583, C584, C585, C586, C587, C588, C589, C590, C591, C592, C593, C594, C595, C596, C597, C598, C599, C600, C601, C602, C603, C604, C605, C606, C607, C608, C609, C610, C611, C612, C613, C614, C615, C616, C617, C618, C619, C620, C621, C622, C623, C624, C625, C626, C627, C628, C629, C630, C631, C632, C633, C634, C635, C636, C637, C638, C639, C640, C641, C642, C643, C644, C645, C646, C647, C648, C649, C650, C651, C652, C653, C654, C655, C656, C657, C658, C659, C660, C661, C662, C663, C664, C665, C666, C667, C668, C669, C670, C671, C672, C673, C674, C675, C676, C677, C678, C679, C680, C681, C682, C683, C684, C685, C686, C687, C688, C689, C690, C691, C692, C693, C694, C695, C696, C697, C698, C699, C700, C701, C702, C703, C704, C705, C706, C707, C708, C709, C710, C711, C712, C713, C714, C715, C716, C717, C718, C719, C720, C721, C722, C723, C724, C725, C726, C727, C728, C729, C730, C731, C732, C733, C734, C735, C736, C737, C738, C739, C740, C741, C742, C743, C744, C745, C746, C747, C748, C749, C750, C751, C752, C753, C754, C755, C756, C757, C758, C759, C760, C761, C762, C763, C764, C765, C766, C767, C768, C769, C770, C771, C772, C773, C774, C775, C776, C777, C778, C779, C780, C781, C782, C783, C784, C785, C786, C787, C788, C789, C790, C791, C792, C793, C794, C795, C796, C797, C798, C799, C800, C801, C802, C803, C804, C805, C806, C807, C808, C809, C810, C811, C812, C813, C814, C815, C816, C817, C818, C819, C820, C821, C822, C823, C824, C825, C826, C827, C828, C829, C830, C831, C832, C833, C834, C835, C836, C837, C838, C839, C840, C841, C842, C843, C844, C845, C846, C847, C848, C849, C850, C851, C852, C853, C854, C855, C856, C857, C858, C859, C860, C861, C862, C863, C864, C865, C866, C867, C868, C869, C870, C871, C872, C873, C874, C875, C876, C877, C878, C879, C880, C881, C882, C883, C884, C885, C886, C887, C888, C889, C890, C891, C892, C893, C894, C895, C896, C897, C898, C899, C900, C901, C902, C903, C904, C905, C906, C907, C908, C909, C910, C911, C912, C913, C914, C915, C916, C917, C918, C919, C920, C921, C922, C923, C924, C925, C926, C927, C928, C929, C930, C931, C932, C933, C934, C935, C936, C937, C938, C939, C940, C941, C942, C943, C944, C945, C946, C947, C948, C949, C950, C951, C952, C953, C954, C955, C956, C957, C958, C959, C960, C961, C962, C963, C964, C965, C966, C967, C968, C969, C970, C971, C972, C973, C974, C975, C976, C977, C978, C979, C980, C981, C982, C983, C984, C985, C986, C987, C988, C989, C990, C991, C992, C993, C994, C995, C996, C997, C998, C999, C1000, C1001, C1002, C1003, C1004, C1005, C1006, C1007, C1008, C1009, C1010, C1011, C1012, C1013, C1014, C1015, C1016, C1017, C1018, C1019, C1020, C1021, C1022, C1023, C1024, C1025, C1026, C1027, C1028, C1029, C1030, C1031, C1032, C1033, C1034, C1035, C1036, C1037, C1038, C1039, C1040, C1041, C1042, C1043, C1044, C1045, C1046, C1047, C1048, C1049, C1050, C1051, C1052, C1053, C1054, C1055, C1056, C1057, C1058, C1059, C1060, C1061, C1062, C1063, C1064, C1065, C1066, C1067, C1068, C1069, C1070, C1071, C1072, C1073, C1074, C1075, C1076, C1077, C1078, C1079, C1080, C1081, C1082, C1083, C1084, C1085, C1086, C1087, C1088, C1089, C1090, C1091, C1092, C1093, C1094, C1095, C1096, C1097, C1098, C1099, C1100, C1101, C1102, C1103, C1104, C1105, C1106, C1107, C1108, C1109, C1110, C1111, C1112, C1113, C1114, C1115, C1116, C1117, C1118, C1119, C1120, C1121, C1122, C1123, C1124, C1125, C1126, C1127, C1128, C1129, C1130, C1131, C1132, C1133, C1134, C1135, C1136, C1137, C1138, C1139, C1140, C1141, C1142, C1143, C1144, C1145, C1146, C1147, C1148, C1149, C1150, C1151, C1152, C1153, C1154, C1155, C1156, C1157, C1158, C1159, C1160, C1161, C1162, C1163, C1164, C1165, C1166, C1167, C1168, C1169, C1170, C1171, C1172, C1173, C1174, C1175, C1176, C1177, C1178, C1179, C1180, C1181, C1182, C1183, C1184, C1185, C1186, C1187, C1188, C1189, C1190, C1191, C1192, C1193, C1194, C1195, C1196, C1197, C1198, C1199, C1200, C1201, C1202, C1203, C1204, C1205, C1206, C1207, C1208, C1209, C1210, C1211, C1212, C1213, C1214, C1215, C1216, C1217, C1218, C1219, C1220, C1221, C1222, C1223, C1224, C1225, C1226, C1227, C1228, C1229, C1230, C1231, C1232, C1233, C1234, C1235, C1236, C1237, C1238, C1239, C1240, C1241, C1242, C1243, C1244, C1245, C1246, C1247, C1248, C1249, C1250, C1251, C1252, C1253, C1254, C1255, C1256, C1257, C1258, C1259, C1260, C1261, C1262, C1263, C1264, C1265, C1266, C1267, C1268, C1269, C1270, C1271, C1272, C1273, C1274, C1275, C1276, C1277, C1278, C1279, C1280, C1281, C1282, C1283, C1284, C1285, C1286, C1287, C1288, C1289, C1290, C1291, C1292, C1293, C1294, C1295, C1296, C1297, C1298, C1299, C1300, C1301, C1302, C1303, C1304, C1305, C1306, C1307, C1308, C1309, C1310, C1311, C1312, C1313, C1314, C1315, C1316, C1317, C1318, C1319, C1320, C1321, C1322, C1323, C1324, C1325, C1326, C1327, C1328, C1329, C1330, C1331, C1332, C1333, C1334, C1335, C1336, C1337, C1338, C1339, C1340, C1341, C1342, C1343, C1344, C1345, C1346, C1347, C1348, C1349, C1350, C1351, C1352, C1353, C1354, C1355, C1356, C1357, C1358, C1359, C1360, C1361, C1362, C1363, C1364, C1365, C1366, C1367, C1368, C1369, C1370, C1371, C1372, C1373, C1374, C1375, C1376, C1377, C1378, C1379, C1380, C1381, C1382, C1383, C1384, C1385, C1386, C1387, C1388, C1389, C1390, C1391, C1392, C1393, C1394, C1395, C1396, C1397, C1398, C1399, C1400, C1401, C1402, C1403, C1404, C1405, C1406, C1407, C1408, C1409, C1410, C1411, C1412, C1413, C1414, C1415, C1416, C1417, C1418, C1419, C1420, C1421, C1422, C1423, C1424, C1425, C1426, C1427, C1428, C1429, C1430, C1431, C1432, C1433, C1434, C1435, C1436, C1437, C1438, C1439, C1440, C1441, C1442, C1443, C1444, C1445, C1446, C1447, C1448, C1449, C1450, C1451, C1452, C1453, C1454, C1455, C1456, C1457, C1458, C1459, C1460, C1461, C1462, C1463, C1464, C1465, C1466, C1467, C1468, C1469, C1470, C1471, C1472, C1473, C1474, C1475, C1476, C1477, C1478, C1479, C1480, C1481, C1482, C1483, C1484, C1485, C1486, C1487, C1488, C1489, C1490, C1491, C1492, C1493, C1494, C1495, C1496, C1497, C1498, C1499, C1500, C1501, C1502, C1503, C1504, C1505, C1506, C1507, C1508, C1509, C1510, C1511, C1512, C1513, C1514, C1515, C1516, C1517, C1518, C1519, C1520, C1521, C1522, C1523, C1524, C1525, C1526, C1527, C1528, C1529, C1530, C1531, C1532, C1533, C1534, C1535, C1536, C1537, C1538, C1539, C1540, C1541, C1542, C1543, C1544, C1545, C1546, C1547, C1548, C1549, C1550, C1551, C1552, C1553, C1554, C1555, C1556, C1557, C1558, C1559, C1560, C1561, C1562, C1563, C1564, C1565, C1566, C1567, C1568, C1569, C1570, C1571, C1572, C1573, C1574, C1575, C1576, C1577, C1578, C1579, C1580, C1581, C1582, C1583, C1584, C1585, C1586, C1587, C1588, C1589, C1590, C1591, C1592, C1593, C1594, C1595, C1596, C1597, C1598, C1599, C1600, C1601, C1602, C1603, C1604, C1605, C1606, C1607, C1608, C1609, C1610, C1611, C1612, C1613, C1614, C1615, C1616, C1617, C1618, C1619, C1620, C1621, C1622, C1623, C1624, C1625, C1626, C1627, C1628, C1629, C1630, C1631, C1632, C1633, C1634, C1635, C1636, C1637, C1638, C1639, C1640, C1641, C1642, C1643, C1644, C1645, C1646, C1647, C1648, C1649, C1650, C1651, C1652, C1653, C1654, C1655, C1656, C1657, C1658, C1659, C1660, C1661, C1662, C1663, C1664, C1665, C1666, C1667, C1668, C1669, C1670, C1671, C1672, C1673, C1674, C1675, C1676, C1677, C1678, C1679, C1680, C1681, C1682, C1683, C1684, C1685, C1686, C1687, C1688, C1689, C1690, C1691, C1692, C1693, C1694, C1695, C1696, C1697, C1698, C1699, C1700, C1701, C1702, C1703, C1704, C1705, C1706, C1707, C1708, C1709, C1710, C1711, C1712, C1713, C1714, C1715, C1716, C1717, C1718, C1719, C1720, C1721, C1722, C1723, C1724, C1725, C1726, C1727, C1728, C1729, C1730, C1731, C1732, C1733, C1734, C1735, C1736, C1737, C1738, C1739, C1740, C1741, C1742, C1743, C1744, C1745, C1746, C1747, C1748, C1749, C1750, C1751, C1752, C1753, C1754, C1755, C1756, C1757, C1758, C1759, C1760, C1761, C1762, C1763, C1764, C1765, C1766, C1767, C1768, C1769, C1770, C1771, C1772, C1773, C1774, C1775, C1776, C1777, C1778, C1779, C1780, C1781, C1782, C1783, C1784, C1785, C1786, C1787, C1788, C1789, C1790, C1791, C1792, C1793, C1794, C1795, C1796, C1797, C1798, C1799, C1800, C1801, C1802, C1803, C1804, C1805, C1806, C1807, C1808, C1809, C1810, C1811, C1812, C1813, C1814, C1815, C1816, C1817, C1818, C1819, C1820, C1821, C1822, C1823, C1824, C1825, C1826, C1827, C1828, C1829, C1830, C1831, C1832, C1

A:Reference number: A19188; MUID:83204002; PMID:6342610
A:Contents: the final paper in a series documenting the sequence, glycosylation site, at
A:Accession: A19188
A:Molecule type: protein
A:Residues: 260-286, 'T', 298-764 <CHR>
R:Campbell, R.D.; Porter, R.R.
Proc. Natl. Acad. Sci. U.S.A. 80, 4464-4468, 1983
A:Title: Molecular cloning and characterization of the gene coding for human complement
A:Reference number: A19947; MUID:83272641; PMID:6308626
A:Accession: A19947
A:Molecule type: DNA
A:Residues: 346-764 <CAM>
A:Cross-references: GB:U00125
A:Accession: B19947
A:Molecule type: mRNA
A:Residues: 339-509 <CA1>
A:Cross-references: GB:U00126; NID:g187723; PIDN:AAA36226.1; PID:g553536
R:Wu, L.; Morley, B.J.; Campbell, R.D.
Cell 48, 331-342, 1987
A:Title: Cell-specific expression of the human complement protein factor B gene: evidenc
A:Reference number: A25971; MUID:87102880; PMID:3643061
A:Accession: B25971
A:Molecule type: DNA
A:Residues: 1-99 <WU>
A:Cross-references: GB:M15082; NID:g187699; PIDN:AAA59625.1; PID:g553534
R:Niemann, M.A.; Brown, A.S.; Miller, E.J.
Biochem. J. 274, 473-480, 1991
A:Title: The principal site of glycation of human complement factor B.
A:Reference number: S14339; MUID:91174758; PMID:2006911
A:Accession: S14339
A:Molecule type: protein
A:Residues: 270-339 <NIE>
A:Note: binding site for carbohydrate to lysine under artificial conditions
R:Morley, B.J.; Campbell, R.D.
EMBO J. 3, 153-157, 1984
A:Title: Internal homologues of the Ba fragment from human complement component factor F
A:Reference number: A44628; MUID:84158524; PMID:6323161
A:Accession: A44628
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 16-225, 'P', 227-259 <MOR>
R:Schwable, W.; Luttig, B.; Sokolowski, T.; Escallier, C.; Weiss, E.H.; Meyer zum Busche
Immunobiology 188, 221-232, 1993
A:Title: Human complement factor B: functional properties of a recombinant zymogen of th
A:Reference number: I54409; MUID:94041399; PMID:8225386
A:Accession: I54409
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-764 <RBS>
A:Cross-references: GB:S67310; NID:g452937; PIDN:AMD13989.1; PID:g4261689
R:Horluchi, T.; Kilm, S.; Matsumoto, M.; Watanabe, I.; Fujita, S.; Volanakis, J.E.
Mol. Immunol. 30, 1587-1592, 1993
A:Title: Human complement factor B: cDNA cloning, nucleotide sequencing, phenotypic conv
A:Reference number: I57824; MUID:94067177; PMID:8247029
A:Accession: I57824
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-31, 'O', 33-764 <RE2>
A:Cross-references: GB:U15702; NID:g291921; PIDN:AAA16820.1; PID:g291922
C:Comment: 292-Cys has a free sulphydryl.
C:Genetics:
A:Gene: GDB:BF
A:Cross-references: GDB:119726; OMIM:138470
A:Map position: 6p21.3-6p21.3
A:Introns: 21/3; 99/3; 346/1; 390/1; 424/1; 470/1; 502/3; 542/1; 593/2; 619/1; 652/3; 65
A:Note: the 11st of introns may be incomplete
C:Complex: complement factor B initially forms an inactive complex with complement facto
ment factor C3b forming active C3/C5 convertase, Ba is released
C:Function:
A:Description: Bb is a serine proteinase; C3/C5 convertase cleaves complement C3 alpha c
A:Pathway: complement alternate pathway

C:Superfamily: complement B/C3; complement factor H repeat homology; trypsin homology; v
C:Keywords: acute phase; complement alternate pathway; duplication; glycoprotein; hydrol
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-764/Product: complement factor B #status experimental <MAT>
F:26-259/Product: complement factor B fragment #status experimental <BAF>
F:37-98/Domain: complement factor H repeat homology <FH1>
F:103-158/Domain: complement factor H repeat homology <FH2>
F:165-218/Domain: complement factor H repeat homology <FH3>
F:260-764/Product: C3/C5 convertase Bb fragment #status experimental <BBF>
F:268-458/Domain: von Willebrand factor type A repeat homology <VFA>
F:482-752/Domain: trypsin homology #status atypical <TRY>
F:37-76, 62-98, 103-145, 131-158, 165-205, 191-218, 478-596, 511-527, 599-615, 656-682, 695-725/D
F:122, 142, 285, 378/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:259-260/Cleavage site: Arg-Lys (complement factor D) #status experimental
F:526, 576, 699/Active site: His, Asp, Ser #status experimental
Query Match 7.4%; Score 113.5; DB 1; Length 764;
Best Local Similarity 19.6%; Pred. No. 0.37; Mismatches 109; Indels 99; Gaps 17;
Matches 66; Conservative 63;
QY 2 OGGRREDG-GPA-----CYGFPDLYFIDKSGSV-----LHNMNIEYFVEQLAH 45
DB 243 EGVDAEDGHGPRGQKRTVLDPSSGSMITYVLDSGSDISGASFTGAKKCLVLLIEKVAS 302
QY 46 KFIISPOLRMSFVFSRGTTLTKLTF---DREQIRQGLEEL---QKVLPGSDTYMHG 97
DB 303 YGVKP--RYGLVYATYATPKIWMVSEADSSNADWVVKQINIEYEDHKLKSGTNT----- 355
QY 98 FEASRQIYENR-----QGY-RTASVITLNTDG-----ELHEDLFFPYSE 136
DB 356 -KRALDAVYSWMSWPDVPEEGMNRTRHVIILMTDGLHMGDPIVYDEIRDLIYKGD 414
QY 137 REANRSRDIGAIYCVG--VKDFNETOLARIADSKOHVPEVNDFOALQGIHSLIKKSC 194
DB 415 RKNRPREDYLDVYFVGPLVNOVINALASKKQNEGVFVKVD--MENTLBDVFPQWIDES- 472
QY 195 IEILAAPEPTTCAGBSFQVVRNGRPHARNDVILCSFKINDSVTLNEKPEFVEDTYLL 254
DB 473 -----QSLSLC-----GMVWEHRKGTD-----YHKPQWAKISV-- 501
QY 255 CPAPILKEVMKALQVNMNDGLSFSSVITTTGC 291
DB 502 ----IRPSKHESCMG-----AVVSEYFVLTAHC 527
RESULT 15
A54849
collagen alpha 1(VII) chain precursor - human
N:Alternate names: procollagen alpha 1(VII) chain
C:Species: Homo sapiens (man)
C:Date: 04-Nov-1994 #sequence revision 04-Nov-1994 #text change 09-Jul-2004
A:Accession: A54849; PH0844; S16316; A50296; I84686
R:Christiano, A.M.; Greenspan, D.S.; Lee, S.; Vitto, J.
J. Biol. Chem. 269, 20256-20262, 1994
A:Title: Cloning of human type VII collagen. Complete primary sequence of the alpha1(VII)
A:Reference number: A54849; MUID:94327588; PMID:8051117
A:Accession: A54849
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-2944 <CHR>
A:Cross-references: UNIPROT:Q02388; GB:L02870; NID:g987124; PIDN:AAA75438.1; PID:g987125
R:Tanaka, T.; Takahashi, K.; Futukawa, F.; Imanura, S.
Biochem. Biophys. Res. Commun. 183, 958-963, 1992
A:Title: Molecular cloning and characterization of type VII collagen cDNA.
A:Reference number: PH0844; MUID:92231902; PMID:1567409
A:Accession: PH0844
A:Molecule type: mRNA
A:Residues: 'EPR', 340-475, 'RALSTASHSTLCRARTKRWHCNCGSHWTRACPCNRPASHRAARAG', 524-528, 'C',
A:Cross-references: DDBJ:D11152; DDBJ:D13694; NID:g453698; PIDN:BA020853.1; PID:g453699
A:Experimental source: keratinocyte
A:Note: the authors translated the codon ACC for residues 394 and 397 as Tyr
R:Parente, M.G.; Chung, L.C.; Rymaenen, J.; Woodley, D.T.; Wynn, K.C.; Bauer, E.A.; Mat
Proc. Natl. Acad. Sci. U.S.A. 88, 6931-6935, 1991

A/Title: Human type VII collagen: cDNA cloning and chromosomal mapping of the gene.
 A/Reference number: S16316; MUID:91334380; PMID:1871109
 A/Accession: S16316
 A/Molecule type: mRNA
 A/Residues: 815-892 'E', 894-1439 <PAR>
 A/Cross-references: GB:M55158; GB:S49017; NID:g180914; PIDN:AAA96439.1; PID:g180915
 A/Experimental source: Keratinocyte
 A/Gammon, W.R.; Abernethy, M.L.; Padilla, K.M.; Priyaanah, P.S.; Cook, M.E.; Wright, J.;
 J. Invest. Dermatol. 99, 691-696, 1992
 A/Title: Noncollagenous (NC1) domain of collagen VII resembles multidomain adhesion prot
 A/Reference number: S16328; MUID:93107742; PMID:1465284
 A/Accession: S16328
 A/Status: translated from GB/EMBL/DBU
 A/Molecule type: mRNA
 A/Residues: 'EPR', 372-517, 'DV', 520-540, 'W', 542-1255 <RES>
 A/Cross-references: GB:S51236; NID:G262308; PIDN:AA24637.1; PID:G262309
 A/Seltzer, J.L.; Eisen, A.Z.; Bauer, E.A.; Morris, N.P.; Gnanville, R.W.; Burgess, R.E.
 J. Biol. Chem. 264, 3822-3826, 1989
 A/Title: Cleavage of type VII collagen by interstitial collagenase and type IV collagen
 A/Reference number: A30296; MUID:89139437; PMID:2537292
 A/Accession: A30296
 A/Molecule type: Protein
 A/Residues: 'A', 1240-1246, 'G', 1248-1250, 'XE', 1253-1255, 'Q', 1257, 'E', 12032, 'C', 2034-2041, '
 A/Note: two reported peptides cannot be reliably located
 R/Greenpan, D.S.
 Hum. Mol. Genet. 2, 273-278, 1993
 A/Title: The carboxyl-terminal half of type VII collagen, including the non-collagenous
 A/Reference number: I48103; MUID:93271985; PMID:8499916
 A/Accession: I48103
 A/Status: preliminary; translated from GB/EMBL/DBU
 A/Molecule type: mRNA
 A/Residues: 2395-2871, 'S', 2873-2944 <RES>
 A/Cross-references: GB:L06862; NID:G388713; PIDN:AAA89196.1; PID:G388714
 R/Christiano, A.M.; Rymaszewski, M.; Utter, J.
 Proc. Natl. Acad. Sci. U.S.A. 91, 3549-3553, 1994
 A/Title: Dominant dystrophic epidermolysis bullosa: identification of a Gly ---> Ser sub
 A/Reference number: A55255; MUID:94224777; PMID:8170945
 A/Contents: annotation
 A/Comment: Prolines and lysines at the third position of the tripeptide repeating unit
 ed and subsequently O-glycosylated.
 C/Genetics:
 A/Gene: GDB:COL7A1; EBR1; EBD1; EB
 A/Cross-references: GDB:128750; OMIM:120120
 A/Map position: 3p21.3-3p21.3
 A/Note: defects in this gene can result in dominant and recessive dystrophic epidermolys
 A/Note: there are 118 introns
 C/Complex: type VII collagen is probably a homotrimer
 C/Function:
 A/Description: structural component of extracellular polymer associated with anchoring f
 C/Keywords: coll; signal sequence #status predicted <SIG>
 F/1-16/Domain: signal sequence #status predicted <SIG>
 F/17-2944/Product: collagen alpha 1(VII) chain #status predicted <MAT>
 F/17-1253/Domain: amino-terminal nonhelical #status predicted <NC1>
 F/36-201/Domain: von Willebrand factor type A repeat homology <VWA1>
 F/231-318/Domain: fibronectin type III repeat homology <FN1>
 F/337-413/Domain: fibronectin type III repeat homology <FN2>
 F/414-502/Domain: fibronectin type III repeat homology <FN3>
 F/508-593/Domain: fibronectin type III repeat homology <FN4>
 F/598-683/Domain: fibronectin type III repeat homology <FN5>
 F/686-771/Domain: fibronectin type III repeat homology <FN6>
 F/776-862/Domain: fibronectin type III repeat homology <FN7>
 F/864-952/Domain: fibronectin type III repeat homology <FN8>
 F/954-1045/Domain: fibronectin type III repeat homology <FN9>
 F/1052-1219/Domain: von Willebrand factor type A repeat homology <VWA2>
 F/1170-1172/Domain: cell attachment (R-G-D) motif
 F/1189-1253/Domain: cytochrome/retinol-rich
 F/1254-2783/Region: interrupted helical
 F/1334-1336/Region: cell attachment (R-G-D) motif
 F/2008-2010/Region: cell attachment (R-G-D) motif
 F/2553-2555/Region: cell attachment (R-G-D) motif
 F/2784-2944/Domain: carboxyl-terminal nonhelical #status predicted <NC2>
 F/2876-2929/Domain: animal Kunitz-type proteinase inhibitor homology <BPI>
 F/337,786,1109/Binding site: carbohydrate (Aan) (covalent) #status predicted

F/2167,2176,2185,2188,2664,2667,2673/Modified site: 4-hydroxyproline (Pro) #status exper
 F/2625,2631/Modified site: 5-hydroxylysine (Lys) #status experimental
 F/2625,2631/Binding site: carbohydrate (Lys) (covalent) #status experimental
 F/2634,2802,2804/Dissulfide bonds: interchain #status predicted

Query Match 7.3%; Score 111.5; DB 2; Length 2944;
 Best Local Similarity 24.1%; Pred. No. 3.2;
 Matches 52; Conservative 42; Mismatches 85; Indels 37; Gaps 11;
 QY 18 DLYFTLKSQSV-LHNNNEIYFVEOLAHNF---ISPO-LRMSFIVSTGTT---LMKL 69
 DB 38 DIVFTLQSSSISGRNFRVRSFLEGLVLPFSGAAGVRFATVQYSDPRTFGLDAL 97
 QY 70 TEDREQIROGLEBLQKVLPGSDTYMHGFERASFOIYENRQGYRTASVILATDGEIHE 129
 DB 98 GSGGDVIR-AIRELS--YKGNRTRTGAALLVHADHVLPLQARAGVAKVCLILIDGK-SQ 153
 QY 130 DLFYSREANRSDLAIIYCVGKOPNETOLARIID--SKDHVPVNDGFOALQGIH 187
 DB 154 DLV---DTAAQRLLKGGQVKLFAVGIKNADPEELKRVASQPTSDFFPVND-FSILRTLLP 209
 QY 188 SILKKSGL-----EILAEPSST 204
 DB 210 LVSRVCTTAGGVPTVRPPDDSTAPRDLVISEPS 245

Search completed: June 13, 2005, 20:06:35
 Job time : 23.1647 secs

THIS PAGE BLANK (USPTO)

XX 13-JUN-2002.
 PD 03-OCT-2001; 2001WO-US030941.
 PF 05-DEC-2000; 2000US-0251481P.
 XX (WISC) WISCONSIN ALUMNI RES FOUND.
 PA Young JAT, Bradley KA, Collier RJ, Mogridge JS;
 PI WPI: 2002-713235/77.
 XX N-PSDB; ABV73881.
 DR Novel isolated polypeptide useful for identifying agent that prevents or
 PT reduces effect of anthrax toxin on host cell, for treating human or non-
 PT human animal suffering from anthrax.
 XX Claim 1, Page 29-30; 45BP; English.
 PS The present sequence is the protein sequence of a human surface-bound
 CC anthrax toxin receptor (ATR), as predicted from an isolated cDNA clone.
 CC Anthrax toxin protective antigen (PA) binds to the ATR at a von
 CC Willebrand factor A domain located in the extracellular domain of ATR.
 CC The invention provides ATR polypeptides and polynucleotides, vectors,
 CC host cells, and transgenic and knock-out animals. It also provides
 CC methods for identifying molecules that bind the ATR and which reduce the
 CC toxicity of anthrax toxin. A claimed method for treating anthrax in a
 CC human or animal involves administering an agent that inhibits binding
 CC between PA and ATR at a level effective to reduce the severity of
 CC anthrax. Suitable agents include ATR or a PA-binding fragment of ATR, a
 CC PA-binding polypeptide at least 80% identical to these, a fusion protein,
 CC a monoclonal or polyclonal antibody, a polysaccharide, a lipid or a
 CC nucleic acid
 XX Sequence 368 AA:
 SQ
 Query Match 100.0%; Score 1526; DB 5; Length 368;
 Best Local Similarity 100.0%; Pred. No. 1.5e-155;
 Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QCGRRDGGPACGCGFDLYFLIDSGSVLHHMNEYVVEQLAKFTSPQIRMFPIVST 60
 DB 28 QCGRRDGGPACGCGFDLYFLIDSGSVLHHMNEYVVEQLAKFTSPQIRMFPIVST 87
 QY 61 RGTILMKLTERRDQIRQGLBELQKVLPGCDTYMEHGFERRASEQIYENRQGYRTASV11A 120
 DB 88 RGTILMKLTERRDQIRQGLBELQKVLPGCDTYMEHGFERRASEQIYENRQGYRTASV11A 147
 QY 121 LTDEGLHEDLFFYSERREANRSRDIGAIYCVGVDFNETQIARIADSKDHVPVNDGFOA 180
 DB 148 LTDEGLHEDLFFYSERREANRSRDIGAIYCVGVDFNETQIARIADSKDHVPVNDGFOA 207
 QY 181 LOGIITHILTKKSCIEIIAAEPSTTCAGSRFQVNVVNGNFRARAVNDRLCSFKINDSVTL 240
 DB 208 LOGIITHILTKKSCIEIIAAEPSTTCAGSRFQVNVVNGNFRARAVNDRLCSFKINDSVTL 267
 QY 241 NEKFPFVEDTYLCPAPILKEVGKMAALQVSMNDGLSISSVITTTTHCSDG 293
 DB 268 NEKFPFVEDTYLCPAPILKEVGKMAALQVSMNDGLSISSVITTTTHCSDG 320

RESULT 2
 ID AAE01439 standard; protein; 403 AA.
 XX AAE01439;
 AC AAE01439;
 XX 17-JUL-2001 (first entry)
 DT Human gene 4 encoded secreted protein HMLFR02, SEQ ID NO:94.
 DE Human; secreted protein; proliferative disorder; cancer; tumour;

KW foetal abnormality; developmental abnormality; haematopoietic disorder;
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
 KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
 KW cardiovascular disorder; angiogenic disorder; kidney disorder;
 KW gastrointestinal disorder; pregnancy-related disorder;
 KW endocrine disorder; infection; wound healing; vulnerability; cell culture;
 KW chemotaxis; food additive; gene therapy; binding partner identification;
 KW chromosome 19.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH Peptide 1..27
 FT /label= signal_peptide
 FT Protein 28..403
 FT /note= "Mature human secreted protein"
 PN W0200134626-A1.
 PD 17-MAY-2001.
 XX 01-NOV-2000; 2000WO-US030045.
 PF 05-NOV-1999; 99US-0163581P.
 PR 30-JUN-2000; 2000US-0215133P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Ruden SM, Komatsoulis GA, Moore PA, Birse CE, Nt J;
 PI WPI: 2001-308778/32.
 DR N-PSDB; AAD05303.
 XX New nucleic acid molecules encoding 28 human secreted proteins for
 PT diagnosing, preventing, treating or ameliorating medical conditions and
 PT used as food additives or preservatives.
 XX Claim 11, Page 485-486; 562BP; English.
 PS
 XX AAD05300-AAD05319 represent cDNAs corresponding to 28 human secreted
 CC protein genes, and AAE01436-AAE01513 represent the proteins they encode.
 CC AAE01514-AAE01544 represent human secreted protein fragments or variants.
 CC The genes and their secreted proteins are useful for preventing, treating
 CC or ameliorating medical conditions, e.g., by protein or gene therapy.
 CC Pathological conditions can be diagnosed by determining the amount of the
 CC new protein in a sample or by determining the presence of mutations in
 CC the new genes. Specific uses are described for each of the 28 genes,
 CC based on the tissues in which they are most highly expressed, and include
 CC developing products for the diagnosis or treatment of proliferative
 CC disorders, cancer, tumours, foetal and developmental abnormalities,
 CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
 CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,
 CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
 CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
 CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
 CC angiogenic disorders, kidney disorders, gastrointestinal disorders,
 CC pregnancy-related disorders, endocrine disorders, and infections. The
 CC proteins can also be used to aid wound healing and epithelial cell
 CC proliferation, to prevent skin aging due to sunburn, to maintain organs
 CC before transplantation, for supporting cell culture of primary tissues,
 CC to regenerate tissues, to identify their cognate ligands or binding
 CC partners, and in chemotaxis, and can be used as a food additive or
 CC preservative to modify storage properties. Antibodies specific for a
 CC protein of the invention can be used in alleviating symptoms associated
 CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,
 CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The
 CC present sequence represents a human secreted protein of the invention
 XX
 SQ Sequence 403 AA;
 Query Match 100.0%; Score 1526; DB 4; Length 403;

Best Local Similarity 100.0%; Pred. No. 1.8e-155; Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGGRREDGPGACGFDLYFLDKSGSVLHHMNEIYFVQOLAKHFTSPQLRMSFIYFST 60
 Db 28 QGGRREDGPGACGFDLYFLDKSGSVLHHMNEIYFVQOLAKHFTSPQLRMSFIYFST 87
 QY 61 RGTTLMLKLTEDREQIRQGLBELQKVLPGGDTYMHGFERASEQIYYENRQGYRTASVITA 120
 Db 88 RGTTLMLKLTEDREQIRQGLBELQKVLPGGDTYMHGFERASEQIYYENRQGYRTASVITA 147
 QY 121 LTDELHEDLFFYSERERANSRDIGATVYCVGVDFNETQLARLADSKDHVPVNDGFQA 180
 Db 148 LTDELHEDLFFYSERERANSRDIGATVYCVGVDFNETQLARLADSKDHVPVNDGFQA 207
 QY 181 LQGIHSLKSKCEIILAAEPSTICAGESFOVVVRGNGFRRARVNDVLCSEFKINDSVTL 240
 Db 208 LQGIHSLKSKCEIILAAEPSTICAGESFOVVVRGNGFRRARVNDVLCSEFKINDSVTL 267
 QY 241 NEKPSVEDTYLLCPAPILKEVGKKAALQVSMNDGLSFISSVIITTHCSG 293
 Db 268 NEKPSVEDTYLLCPAPILKEVGKKAALQVSMNDGLSFISSVIITTHCSG 320

RESULT 3
 ABG63874 ID ABG63874 standard; protein; 403 AA.
 XX AC ABG63874;
 XX DT 27-AUG-2002 (first entry)
 XX DE Human albumin fusion protein #549.
 XX KM Albumin fusion protein; therapeutic protein X; human albumin; HA;
 KM human serum albumin; HSA; cancer; reproductive disorder;
 KM digestive disorder; immune disorder; endocrine disorder;
 KM haematopoietic disorder; neural disorder; connective disorder;
 KM cytoskeletal; antiinfectility; antiinflammatory; anticancer;
 KM immunomodulator; anti-HIV; antidiabetic; haemostatic; noctropic;
 KM neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
 KM osteopathic; antichratic.
 XX OS Homo sapiens.
 XX OS Synthetic.
 XX PN MO200177137-A1.
 XX PD 18-OCT-2001.
 XX PF 12-APR-2001; 2001WO-US011988.
 XX PR 12-APR-2000; 2000US-0229358P.
 XX PR 25-APR-2000; 2000US-019384P.
 XX PR 21-DEC-2000; 2000US-0256931P.
 XX PA (HUMA-) HUMAN GENOME SCI INC.
 XX PI Rosen CA, Haseltine WA;
 XX DR WPI; 2002-010886/01.
 XX PT New fusion protein for treating disease e.g. diabetes comprises an
 XX albumin fused to a therapeutic protein.
 XX PS Claim 1, Page 874-875, 2102pp; English.
 XX CC The present invention relates to albumin fusion proteins comprising a
 CC therapeutic protein X and human albumin (HA), also known as human serum
 CC albumin, HSA). The proteins are useful for treating a disease or disorder
 CC that may be modulated by therapeutic protein X. The albumin extends the
 CC shelf-life of protein X, and may increase its biological in vitro/in vivo
 CC activity. The protein is useful for treating and diagnosing disorders

CC such as cancer, reproductive disorders, digestive disorders (e.g. Crohn's
 CC disease, ulcerative colitis), immune disorders (e.g. acquired
 CC immunodeficiency syndrome, AIDS), endocrine disorders (e.g. diabetes),
 CC haematopoietic disorders, neural disorders (e.g. Alzheimer's,
 CC Parkinson's, Creutzfeldt-Jacob disease, encephalomyelitis, meningitis,
 CC schizophrenia), and connective disorders (e.g. osteoporosis, arthritis).
 CC ABG63326-ABG65518 represent albumin fusion proteins of the invention
 XX
 SQ Sequence 403 AA;

Query Match 100.0%; Score 1526; DB 5; Length 403;
 Best Local Similarity 100.0%; Pred. No. 1.8e-155; Indels 0; Gaps 0;
 Matches 293; Conservative 0; Mismatches 0;

QY 1 QGGRREDGPGACGFDLYFLDKSGSVLHHMNEIYFVQOLAKHFTSPQLRMSFIYFST 60
 Db 28 QGGRREDGPGACGFDLYFLDKSGSVLHHMNEIYFVQOLAKHFTSPQLRMSFIYFST 87
 QY 61 RGTTLMLKLTEDREQIRQGLBELQKVLPGGDTYMHGFERASEQIYYENRQGYRTASVITA 120
 Db 88 RGTTLMLKLTEDREQIRQGLBELQKVLPGGDTYMHGFERASEQIYYENRQGYRTASVITA 147
 QY 121 LTDELHEDLFFYSERERANSRDIGATVYCVGVDFNETQLARLADSKDHVPVNDGFQA 180
 Db 148 LTDELHEDLFFYSERERANSRDIGATVYCVGVDFNETQLARLADSKDHVPVNDGFQA 207
 QY 181 LQGIHSLKSKCEIILAAEPSTICAGESFOVVVRGNGFRRARVNDVLCSEFKINDSVTL 240
 Db 208 LQGIHSLKSKCEIILAAEPSTICAGESFOVVVRGNGFRRARVNDVLCSEFKINDSVTL 267
 QY 241 NEKPSVEDTYLLCPAPILKEVGKKAALQVSMNDGLSFISSVIITTHCSG 293
 Db 268 NEKPSVEDTYLLCPAPILKEVGKKAALQVSMNDGLSFISSVIITTHCSG 320

RESULT 4
 ADL77139 ID ADL77139 standard; protein; 403 AA.
 XX AC ADL77139;
 XX DT 20-MAY-2004 (first entry)
 XX DE Albumin fusion protein related therapeutic protein X, SEQ ID No 621.
 XX KM albumin fusion protein; cytostatic; antianaemic; antiarthritis;
 KM antiaesthetic; anti-HIV; immunosuppressive; antiinflammatory;
 KM antipsoaratic; antibacterial; osteopathic; dermatological; antigout;
 KM immunomodulator; antiarhythmic; cardiant; noctropic; antilipemic;
 KM nephrotropic; utropathic; neuroprotective; antiparkinsonian; tranquilizer;
 KM antidiabetic; anabolic; hypertensive; vulnery; gene therapy; cancer;
 KM reproductive system disorder; therapeutic protein.
 XX OS Unidentified.
 XX OS US2004010134-A1.
 XX PN 15-JAN-2004.
 XX PD 12-APR-2001; 2001US-00833245.
 XX PR 12-APR-2000; 2000US-0229358P.
 XX PR 25-APR-2000; 2000US-019384P.
 XX PR 21-DEC-2000; 2000US-0256931P.
 XX PA (ROSE/) ROSEN C A.
 XX PA (HASE/) HASELTINE W A.
 XX PI Rosen CA, Haseltine WA;
 XX DR WPI; 2004-090519/09.
 XX PT New albumin fusion proteins, useful for diagnosing, creating, preventing

PT or ameliorating diseases or disorders e.g. cancer, anemia, arthritis,
 PT asthma, inflammatory bowel disease or Alzheimer's disease.
 XX
 PS Disclosure; SEQ ID NO 621; 279bp; English.
 XX
 CC The invention relates to a novel albumin fusion protein. The invention
 CC further relates to: a composition comprising the albumin fusion protein
 CC and a pharmaceutical carrier; a kit comprising the composition of the
 CC albumin fusion protein formula; a method of treating a disease or
 CC disorder in a patient comprising the step of administering the albumin
 CC fusion protein; a method of treating a patient with a disease or disorder
 CC that is modulated by therapeutic protein: X, or its fragment or variant;
 CC a method of extending the shelf life of Therapeutic protein: X, or its
 CC fragment or variant; a nucleic acid molecule comprising a polynucleotide
 CC sequence encoding the albumin fusion protein; a vector comprising the
 CC nucleic acid molecule of the albumin fusion protein; and a host cell
 CC comprising the nucleic acid molecule of the albumin fusion protein. The
 CC albumin fusion protein and its compositions have the following
 CC activities: cytostatic, antineoplastic, antitumor, antidiabetic, anti-
 CC HIV, immunosuppressive, antiinflammatory, antipruritic, antibacterial,
 CC osteoprotic, dermatological, antipain, immunomodulator, antiallergic,
 CC cardiatic, neuroprotective, antiparkinsonian, tranquilizer, uropathic,
 CC neuroprotective, and antipruritic. The albumin fusion protein nucleic acid may
 CC be used in gene therapy to treat disorders. The albumin fusion protein is
 CC useful for diagnosing, treating, preventing or ameliorating diseases or
 CC disorders comprising indication: Y. The diseases or disorders include:
 CC cancer (e.g. leukaemia, colon, bone, breast, liver or lung cancer),
 CC immune or haematopoietic diseases (e.g. anaemia, Hodgkin's disease, acute
 CC lymphocytic anaemia, multiple myeloma, arthritis, asthma, AIDS,
 CC autoimmune disease, inflammatory bowel disease, psoriasis or Lyme
 CC disease), reproductive system disorders (e.g. prostaticitis, inguinal
 CC hernia, varicocele, penile carcinoma, ovarian adenocarcinoma or Sertoli-
 CC Leydig tumours), musculoskeletal diseases (e.g. giant cell tumours,
 CC Paget's disease, systemic lupus erythematosus, gout, muscular dystrophy
 CC or cachexia), cardiovascular disease (e.g. rhodomyoma, heart disease,
 CC arrhythmia, cardiac arrest, heart valve disease, hypernatraemia or
 CC hyponatremia), mixed foetal diseases (e.g. foetal alcohol syndrome,
 CC Down's syndrome, Patou syndrome, Turner's syndrome, Apert syndrome or Tay-
 CC -Sachs disease), excretory diseases (e.g. urinary incontinence, urinary
 CC tract infections or renal disorders), neural or sensory disease (e.g.
 CC Alzheimer's disease, Parkinson's disease, cerebral malaria, meningitis,
 CC cerebellar ataxia, attention deficit disorder, autism or obsessive
 CC compulsive disorder), respiratory disease (e.g. emphysema, lung cancer or
 CC occupational lung disease), endocrine diseases (e.g. diabetes, Addison's
 CC disease or glomerulonephritis), digestive diseases (e.g. portal
 CC hypertension, irritable bowel disease, gastric atrophy or pancreatitis)
 CC or connective tissue or epithelial diseases (e.g. Crohn's disease,
 CC scleroderma, wound healing or epidermolysis bullosa). This sequence
 CC represents a therapeutic protein X relating to the albumin fusion protein
 CC of the invention. The sequence listing data for this specification was
 CC downloaded from the USPTO website.
 CC
 XX
 XX
 SQ Sequence 403 AA;
 Query Match 100.0%; Score 1526; DB 8; Length 403;
 Best Local Similarity 100.0%; Pred. No. 1.8e-155;
 Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QGGRREDGPGACGDFDLYFLDKSGSVLHNMNEIYFVVEQLAKHFIQRLMSFIVST 60
 DB 28 QGGRREDGPGACGDFDLYFLDKSGSVLHNMNEIYFVVEQLAKHFIQRLMSFIVST 87
 QY 61 RGTTLMLKLTEDRQIRQGLLELOKVLPGSDTYMHGFERASBOIYYENRGYRTASVIA 120
 DB 88 RGTTLMLKLTEDRQIRQGLLELOKVLPGSDTYMHGFERASBOIYYENRGYRTASVIA 147
 QY 121 LTGDELHEDLFFYSERANRSDLGAIYVCVGDQFNEFOTLARIADSKDHVPVNDGFOA 180
 DB 148 LTGDELHEDLFFYSERANRSDLGAIYVCVGDQFNEFOTLARIADSKDHVPVNDGFOA 207
 QY 161 LOGIHSILKKSCEIILAEPSSTCAGRSPOVYVNGNFRHARNDRVLCSEFKINDSVTL 240

DB 208 LOGIHSILKKSCEIILAEPSSTCAGRSPOVYVNGNFRHARNDRVLCSEFKINDSVTL 267
 QY 241 NEKPFSEVEDTYLLCPAPILKEVGKMAALQVSNNDGSEFISSVITTTTHCSGD 293
 DB 268 NEKPFSEVEDTYLLCPAPILKEVGKMAALQVSNNDGSEFISSVITTTTHCSGD 320
 RESULT 5
 AD100550
 ID AD100550 standard; protein; 551 AA.
 XX
 AC AD100550;
 XX
 DT 22-APR-2004 (first entry)
 XX
 XX Human TANGO 197 Ig fusion mutated protein - plasmid p0610.
 XX
 KM fusion; von Willebrand factor A-like domain; vWF; antibacterial;
 KM cutaneous; inhalation anthrax; human; TANGO 197 Ig fusion; mutant;
 KM plasmid p0610; mutain.
 XX
 OS Homo sapiens.
 XX Synthetic.
 XX US2003144193-A1.
 XX 31-JUL-2003.
 XX 24-JUL-2002; 2002US-00201292.
 XX 20-DEC-2001; 2001US-00038307.
 XX
 PA (ROTT/) ROTTMAN J B.
 PA (OKEE/) O'KEEFE T L.
 PA (OZKA/) OZKAYNAK E.
 PA (HEAL/) HEALEY J J.
 XX
 PT Rottman JB, O'keefe TL, Ozkaynak E, Healey JJ;
 XX
 DR WPI; 2003-720708/68.
 DR N-PSDB; AD100549.
 XX
 PT New TANGO 197 and/or TANGO 216 fusion polypeptides for preventing or
 PT ameliorating symptoms of anthrax comprises a von Willebrand factor A-like
 PT domain (vWF) amino acid sequence and an amino acid sequence heterologous
 PT to the vWF.
 PS Claim 45; SEQ ID NO 18; 86pp; English.
 XX
 XX The invention relates to a novel fusion polypeptide comprising a von
 CC Willebrand factor A-like domain (vWF) amino acid sequence and an amino
 CC acid sequence heterologous to the vWF. The polypeptide of the invention
 CC demonstrates antibacterial activities whilst the composition and method
 CC may be useful in preventing or ameliorating the symptoms of cutaneous
 CC and/or inhalation anthrax. The current sequence is that of the human
 CC TANGO 197 Ig mutated fusion protein of the invention.
 CC
 XX
 XX
 SQ Sequence 551 AA;
 Query Match 100.0%; Score 1526; DB 7; Length 551;
 Best Local Similarity 100.0%; Pred. No. 2.8e-155;
 Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QGGRREDGPGACGDFDLYFLDKSGSVLHNMNEIYFVVEQLAKHFIQRLMSFIVST 60
 DB 28 QGGRREDGPGACGDFDLYFLDKSGSVLHNMNEIYFVVEQLAKHFIQRLMSFIVST 87
 QY 61 RGTTLMLKLTEDRQIRQGLLELOKVLPGSDTYMHGFERASBOIYYENRGYRTASVIA 120
 DB 88 RGTTLMLKLTEDRQIRQGLLELOKVLPGSDTYMHGFERASBOIYYENRGYRTASVIA 147
 QY 121 LTGDELHEDLFFYSERANRSDLGAIYVCVGDQFNEFOTLARIADSKDHVPVNDGFOA 180

Db 148 LTDELHEDLFYFSERANRSDGAIIVYCVGVDFNETQLARIADSKDHVPVNDGFQA 207
Qy 181 LQGIHSILKKSCEIILAEPSTICAGESFOVVVRGNGFRHARVDRVLCSEFKINDSVTL 240
Db 208 LQGIHSILKKSCEIILAEPSTICAGESFOVVVRGNGFRHARVDRVLCSEFKINDSVTL 267
Qy 241 NEKPSVEDTYLLCPAPILKEVGKALQVSMNDGLSFSSSVITTTTHCSGD 293
Db 268 NEKPSVEDTYLLCPAPILKEVGKALQVSMNDGLSFSSSVITTTTHCSGD 320

RESULT 6

ADM64576
ID ADM64576 standard; protein; 551 AA.

AC ADM64576;

XX 03-JUN-2004 (first entry)

XX Mouse TANGO197-immunoglobulin (Ig) fusion protein.

XX antibacterial; gene therapy;

KM von Willebrand factor A-like domain amino acid sequence;

KM vWF amino acid sequence; anthrax; Bacillus anthracis; cutaneous anthrax;

KM inhalation anthrax; mouse; TANGO197; immunoglobulin; Ig; fusion protein;

XX gene.

XX gene.

XX gene.

XX gene.

XX gene.

XX gene.

XX gene.

XX gene.

XX gene.

XX gene.

XX gene.

XX gene.

XX gene.

XX gene.

XX gene.

XX gene.

XX gene.

XX gene.

XX gene.

XX gene.

XX gene.

XX gene.

Qy 1 QGRRRGGGPAICGGFDLYFLIDKSGSVLHWNIEIYFVEQLAHKFTSPQLRMSFIYFST 60
Db 28 QGRRRGGGPAICGGFDLYFLIDKSGSVLHWNIEIYFVEQLAHKFTSPQLRMSFIYFST 87
Qy 61 RGTLMKLTERRRQIRQGLELQVLPQSDTYMHEGGERASQIYVNRQGYRTASVITA 120
Db 88 RGTLMKLTERRRQIRQGLELQVLPQSDTYMHEGGERASQIYVNRQGYRTASVITA 147
Qy 121 LTDELHEDLFYFSERANRSDGAIIVYCVGVDFNETQLARIADSKDHVPVNDGFQA 180
Db 148 LTDELHEDLFYFSERANRSDGAIIVYCVGVDFNETQLARIADSKDHVPVNDGFQA 207
Qy 181 LQGIHSILKKSCEIILAEPSTICAGESFOVVVRGNGFRHARVDRVLCSEFKINDSVTL 240
Db 208 LQGIHSILKKSCEIILAEPSTICAGESFOVVVRGNGFRHARVDRVLCSEFKINDSVTL 267
Qy 241 NEKPSVEDTYLLCPAPILKEVGKALQVSMNDGLSFSSSVITTTTHCSGD 293
Db 268 NEKPSVEDTYLLCPAPILKEVGKALQVSMNDGLSFSSSVITTTTHCSGD 320

RESULT 7

ABB90750
ID ABB90750 standard; protein; 564 AA.

AC ABB90750;

XX 30-MAY-2002 (first entry)

XX Human Tumour Endothelial Marker polypeptide SEQ ID NO 232.

XX Human; mouse; rat; TEM; tumour endothelial marker; TEM; PEM; cytoelastic;

KM normal endothelial marker; pan-endothelial marker; immunostimulant;

KM antiangiogenic; tumour; neovascularisation; vascularised tumour;

KM polygenic kidney disease; diabetes; retinopathy; rheumatoid arthritis;

KM psoriasis.

XX Homo sapiens.

XX WO200210217-A2.

XX 07-FEB-2002.

XX 01-AUG-2001; 2001WO-US024031.

XX 02-AUG-2000; 2000US-0222599P.

XX 11-AUG-2000; 2000US-0224360P.

XX 11-APR-2001; 2001US-0282850P.

XX (UJO) UNIV JOHNS HOPKINS.

XX St Croix B, Kinzler KW, Vogelstein B;

XX WPI; 2002-291856/33.

XX N-PSDB; ABL92104.

XX The invention relates to an isolated molecule comprising an antibody

XX variable region which specifically binds to an extracellular domain of a

XX tumour endothelial marker (TEM) protein selected from ABB90732, ABB90740,

XX ABB90749, ABB90750 and ABB90769. The antibodies which bind to TEM

XX proteins have cytoelastic, immunostimulant and antiangiogenic activity.

XX They are useful for inhibiting tumour growth, neovascularisation in subjects

XX bearing a vascularised tumour, polycystic kidney disease, diabetic

XX retinopathy, rheumatoid arthritis and psoriasis. Human, mouse and rat TEM

XX genes and the encoded proteins (ABL92075-ABL92141 and ABB90721-ABB90789)

XX are disclosed, as are marker oligonucleotide sequences: tumour

XX endothelial markers (TEM) ABL91996-ABL92041 and ABL92143-ABL92191; normal

CC endothelial markers (NEM) AB192042-AB192074; and pan-endothelial markers
 CC (PEM) AB191903-AB191995
 XX
 SQ Sequence 564 AA;

Query Match 100.0%; Score 1526; DB 5; Length 564;
 Best Local Similarity 100.0%; Pred. No. 2,9e-155;
 Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGRRBDGACGCGFDLYFIIDKSSVYLHNMNIYFVEQLAHKFTSPQLRMSFIYEST 60
 DB 28 QGRRBDGACGCGFDLYFIIDKSSVYLHNMNIYFVEQLAHKFTSPQLRMSFIYEST 87

QY 61 RGTTLMLKLTEDREQIRQGLELQKVLPGGDTYHHEGFERASEQIYYENRGYRTASV11A 120
 DB 88 RGTTLMLKLTEDREQIRQGLELQKVLPGGDTYHHEGFERASEQIYYENRGYRTASV11A 147

QY 121 LTDBELHEDLFFYSERBANRSRDGAIVYCVGVKDFNETQLARIADSKDHVPVNDGFOA 180
 DB 148 LTDBELHEDLFFYSERBANRSRDGAIVYCVGVKDFNETQLARIADSKDHVPVNDGFOA 207

QY 181 LOGIHSILKKSCEILAAEPSTICAGESFOVVVRNGGPFHARVNDVLCSPKINDSVTL 240
 DB 208 LOGIHSILKKSCEILAAEPSTICAGESFOVVVRNGGPFHARVNDVLCSPKINDSVTL 267

QY 241 NEKPFVEDTYLLCPAPILKEVGKKAALQVSMNDGLSFISSSVITTTTHCSGD 293
 DB 268 NEKPFVEDTYLLCPAPILKEVGKKAALQVSMNDGLSFISSSVITTTTHCSGD 320

RESULT 8
 ABB90724
 ID ABB90724 standard; protein; 564 AA.

AC ABB90724;
 XX
 DT 30-MAY-2002 (first entry)
 XX
 DE Human Tumour Endothelial Marker polypeptide SEQ ID NO 187.
 XX
 KM Human; mouse; rat; TEM; tumour endothelial marker; NEM; PEM; cytostatic;
 KM normal endothelial marker; pan-endothelial marker; immunostimulant;
 KM antiangiogenic; tumour; neovascularization; vascularised tumour;
 KM polycystic kidney disease; diabetes; reclinopathy; rheumatoid arthritis;
 KM psoriasis.
 XX
 OS Homo sapiens.
 XX
 PN WO200210217-A2.
 XX
 PD 07-FEB-2002.
 XX
 PF 01-AUG-2001; 2001WO-US024031.
 XX
 PR 02-AUG-2000; 2000US-0222599P.
 PR 11-AUG-2000; 2000US-0224360P.
 PR 11-APR-2001; 2001US-0282850P.
 XX
 PA (UUGO) UNIV JOHNS HOPKINS.
 XX
 PI St Croix B, Kinzler KM, Vogelstein B;
 XX
 DR WPI; 2002-291856/33.
 XX
 PT An isolated molecule comprising an antibody variable region which
 PT specifically binds to an extracellular domain of a tumor endothelial
 PT marker (TEM) protein, useful for inhibiting tumor growth.
 XX
 PS Disclosure; Page 136-137; 33pp; English.
 XX
 CC The invention relates to an isolated molecule comprising an antibody
 CC variable region which specifically binds to an extracellular domain of a
 CC tumour endothelial marker (TEM) protein selected from ABB90732, ABB90740,

CC ABB90749, ABB90750 and ABB90769. The antibodies which bind to TEM
 CC proteins have cytostatic, immunostimulant and antiangiogenic activity.
 CC They are useful for inhibiting tumour growth, neovascularization in subjects
 CC bearing a vascularised tumour, polycystic kidney disease, diabetic
 CC retinopathy, rheumatoid arthritis and psoriasis. Human, mouse and rat TEM
 CC genes and the encoded proteins (AB192075-AB192141 and ABB90721-ABB90789)
 CC are disclosed, as are marker oligonucleotide sequences: tumour
 CC endothelial markers (TEM) AB191996-AB192041 and AB192143-AB192191; normal
 CC endothelial markers (NEM) AB192042-AB192074; and pan-endothelial markers
 CC (PEM) AB191903-AB191995
 XX
 SQ Sequence 564 AA;

Query Match 100.0%; Score 1526; DB 5; Length 564;
 Best Local Similarity 100.0%; Pred. No. 2,9e-155;
 Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGRRBDGACGCGFDLYFIIDKSSVYLHNMNIYFVEQLAHKFTSPQLRMSFIYEST 60
 DB 28 QGRRBDGACGCGFDLYFIIDKSSVYLHNMNIYFVEQLAHKFTSPQLRMSFIYEST 87

QY 61 RGTTLMLKLTEDREQIRQGLELQKVLPGGDTYHHEGFERASEQIYYENRGYRTASV11A 120
 DB 88 RGTTLMLKLTEDREQIRQGLELQKVLPGGDTYHHEGFERASEQIYYENRGYRTASV11A 147

QY 121 LTDBELHEDLFFYSERBANRSRDGAIVYCVGVKDFNETQLARIADSKDHVPVNDGFOA 180
 DB 148 LTDBELHEDLFFYSERBANRSRDGAIVYCVGVKDFNETQLARIADSKDHVPVNDGFOA 207

QY 181 LOGIHSILKKSCEILAAEPSTICAGESFOVVVRNGGPFHARVNDVLCSPKINDSVTL 240
 DB 208 LOGIHSILKKSCEILAAEPSTICAGESFOVVVRNGGPFHARVNDVLCSPKINDSVTL 267

QY 241 NEKPFVEDTYLLCPAPILKEVGKKAALQVSMNDGLSFISSSVITTTTHCSGD 293
 DB 268 NEKPFVEDTYLLCPAPILKEVGKKAALQVSMNDGLSFISSSVITTTTHCSGD 320

RESULT 9
 ABB54904
 ID ABB54904 standard; protein; 564 AA.

AC ABB54904;
 XX
 DT 08-JAN-2003 (first entry)
 XX
 DE Human anthrax toxin receptor.
 XX
 KM Anthrax; toxin; receptor; human; TEM8; antibacterial.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FH Peptide 1..27
 FT /label= Signal_peptide
 FT Protein 28..564
 FT /label= Mature_protein
 FT Domain 28..320
 FT /note= "extracellular domain"
 FT Domain 44..216
 FT /note= "von Willebrand factor A domain"
 FT Domain 320..343
 FT /note= "putative transmembrane domain"
 FT Domain 344..564
 FT /note= "cytoplasmic domain"
 XX
 PN WO200246228-A2.
 XX
 PD 13-JUN-2002.
 XX
 PF 03-OCT-2001; 2001WO-US030941.
 XX
 PR 05-DEC-2000; 2000US-0251481P.

XX (WISC) WISCONSIN ALUMNI RES FOUND.
 PA Young JAT, Bradley KA, Collier RJ, Mogridge JS;
 PI WPI; 2002-713235/77.
 XX N-PSDB; ABV73882.
 DR Novel isolated polypeptide useful for identifying agent that prevents or
 XX reduces effect of anthrax toxin on host cell, for treating human or non-
 PT human animal suffering from anthrax.
 PS Claim 1; Page 37-39; 45pp; English.
 XX The present sequence is the protein sequence of a human surface-bound
 CC anthrax toxin receptor (ATR), previously designated TEM8, and identified
 CC following a database screening using a newly isolated human surface-bound
 CC ATR (see ABP54903). The present sequence differs from the newly isolated
 CC human ATR only in the cytoplasmic domain (221 rather than 25 amino acids
 CC long), suggesting differential splicing of a primary mRNA transcript.
 CC Anthrax toxin protective antigen (PA) binds to these ATRs at a von
 CC Willebrand factor A domain located in the extracellular domain of ATR.
 CC The invention provides ATR polypeptides and polynucleotides, vectors,
 CC host cells, and transgenic and knock-out animals. It also provides
 CC methods for identifying molecules that bind the ATR and which reduce the
 CC toxicity of anthrax toxin. A claimed method for treating anthrax in a
 CC human or animal involves administering an agent that inhibits binding
 CC between PA and ATR at a level effective to reduce the severity of
 CC anthrax. Suitable agents include the present polypeptide or a PA-binding
 CC fragment of it, a PA-binding polypeptide at least 80% identical to these,
 CC a fusion protein, a monoclonal or polyclonal antibody, a polysaccharide,
 CC a lipid or a nucleic acid
 XX
 SQ Sequence 564 AA;
 Query Match 100.0%; Score 1526; DB 5; Length 564;
 Best Local Similarity 100.0%; Pred. No. 2.9e-155;
 Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QGGRREDGPGACYGFDLYFLIDKSGSVLHHMNEIYFVEQLAHKFTSPQLRMSFIYFST 60
 Db 28 QGGRREDGPGACYGFDLYFLIDKSGSVLHHMNEIYFVEQLAHKFTSPQLRMSFIYFST 87
 QY 61 RGTTLMLKLTEDREIQROGLELQVLPFGDPTVMHEGFERASEQIYYENRQGYRTASVYIA 120
 Db 88 RGTTLMLKLTEDREIQROGLELQVLPFGDPTVMHEGFERASEQIYYENRQGYRTASVYIA 147
 QY 121 LTDELHEDLFFYSERANRSRDIGAIYCVGVKDFNETQLARIADSKDHPVPVNDGFOA 180
 Db 148 LTDELHEDLFFYSERANRSRDIGAIYCVGVKDFNETQLARIADSKDHPVPVNDGFOA 207
 QY 181 LOGIHSILKSKSCIEILAAEPSTICAGESFOVVVRGNGFPHANVDRVLCSPKINDSVTL 240
 Db 208 LOGIHSILKSKSCIEILAAEPSTICAGESFOVVVRGNGFPHANVDRVLCSPKINDSVTL 267
 QY 241 NEKPFSEVDYLLCPAPILKEVGKKAALQVSMNDGLSFISSVYIITTHCSGD 293
 Db 268 NEKPFSEVDYLLCPAPILKEVGKKAALQVSMNDGLSFISSVYIITTHCSGD 320

RESULT 10
 ABUS4457
 ID ABUS4457 standard; protein; 564 AA.
 XX
 AC ABUS4457;
 XX
 DT 12-MAR-2003 (first entry)
 XX
 DE Human tumour endothelial marker TEM 19.
 XX
 KM Human, endothelial cell; EC; tumour endothelial cell; TEM; NEM;
 KM Tumour endothelial marker; normal endothelial marker; PEM;
 KM pan-endothelial marker; polycystic kidney disease; psoriasis;

KM diabetic retinopathy; rheumatoid arthritis; tumour angiogenesis;
 KM neovascularization; immune response; cytotoxic; antidiabetic;
 KM opthalmological; antineumatic; antiarthritic; antipsoriatic.
 XX Homo sapiens.
 XX WO200283874-A2.
 XX 24-OCT-2002.
 PD
 XX 10-APR-2002; 2002WO-US008253.
 PR 11-APR-2001; 2001US-0282850P.
 PR 06-FEB-2002; 2002US-0354262P.
 XX (UJVO) UNIV JOHNS HOPKINS.
 PA Carson-Walter E, St Croix B, Kinzler KW, Vogelstein B;
 PI WPI; 2003-093016/08.
 DR N-PSDB; ABX72029.
 XX
 PS New purified human transmembrane protein, designated as tumor endothelial
 PT marker (TEM) 3, useful for detecting, diagnosing or treating tumors,
 PT polycystic kidney disease, diabetic retinopathy, rheumatoid arthritis or
 PS psoriasis.
 XX Disclosure; Page 226-227; 374pp; English.
 XX The present invention relates to a novel method for the isolation of
 CC endothelial cells (ECs), and the identification of genes expressed in
 CC normal and tumour ECs. Tumour endothelial marker (TEM), normal
 CC endothelial marker (NEM), and pan-endothelial marker (PEM) genes are
 CC identified in human ECs. The human EC marker proteins and the
 CC polynucleotide sequences encoding them are useful for detecting,
 CC diagnosing or treating tumours as well as polycystic kidney disease,
 CC diabetic retinopathy, rheumatoid arthritis, and psoriasis. They are also
 CC useful for inhibiting neovascularization or tumour angiogenesis, for
 CC inducing an immune response to tumour endothelial cells in a patient, or
 CC for identifying candidate drugs for treating tumours. The present
 CC sequence represents a human TEM or NEM protein of the invention
 XX
 SQ Sequence 564 AA;
 Query Match 100.0%; Score 1526; DB 6; Length 564;
 Best Local Similarity 100.0%; Pred. No. 2.9e-155;
 Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QGGRREDGPGACYGFDLYFLIDKSGSVLHHMNEIYFVEQLAHKFTSPQLRMSFIYFST 60
 Db 28 QGGRREDGPGACYGFDLYFLIDKSGSVLHHMNEIYFVEQLAHKFTSPQLRMSFIYFST 87
 QY 61 RGTTLMLKLTEDREIQROGLELQVLPFGDPTVMHEGFERASEQIYYENRQGYRTASVYIA 120
 Db 88 RGTTLMLKLTEDREIQROGLELQVLPFGDPTVMHEGFERASEQIYYENRQGYRTASVYIA 147
 QY 121 LTDELHEDLFFYSERANRSRDIGAIYCVGVKDFNETQLARIADSKDHPVPVNDGFOA 180
 Db 148 LTDELHEDLFFYSERANRSRDIGAIYCVGVKDFNETQLARIADSKDHPVPVNDGFOA 207
 QY 181 LOGIHSILKSKSCIEILAAEPSTICAGESFOVVVRGNGFPHANVDRVLCSPKINDSVTL 240
 Db 208 LOGIHSILKSKSCIEILAAEPSTICAGESFOVVVRGNGFPHANVDRVLCSPKINDSVTL 267
 QY 241 NEKPFSEVDYLLCPAPILKEVGKKAALQVSMNDGLSFISSVYIITTHCSGD 293
 Db 268 NEKPFSEVDYLLCPAPILKEVGKKAALQVSMNDGLSFISSVYIITTHCSGD 320

RESULT 11
 ABUS4431
 ID ABUS4431 standard; protein; 564 AA.
 XX

AC ABUS4431;
 XX
 DT 12-MAR-2003 (first entry)
 XX
 DE Human tumour endothelial marker TEM 8.
 XX
 KM Human; endothelial cell; EC; tumour endothelial cell; TEM; NEM;
 KM Tumour endothelial marker; normal endothelial marker; PEM;
 KM pan-endothelial marker; polycystic kidney disease; psoriasis;
 KM diabetic retinopathy; rheumatoid arthritis; tumour angiogenesis;
 KM neovascularization; immune response; cytostatic; antidiabetic;
 KM ophthalmological; antineoplastic; antirheumatic; antipsoriatic.
 XX
 OS Homo sapiens.
 XX
 PN WC0200283874-A2.
 XX
 PD 24-OCT-2002.
 XX
 PF 10-APR-2002; 2002WO-US008253.
 XX
 PR 11-APR-2001; 2001US-0282850P.
 PR 06-FEB-2002; 2002US-0354262P.
 XX
 PA (UYJO) UNIV JOHNS HOPKINS.
 XX
 PI Carson-Walter E, St Croix B, Kinzler KM, Vogelstein B;
 XX
 DR WPI; 2003-093016/08.
 XX
 PT New purified human transmembrane protein, designated as tumor endothelial
 PT marker (TEM) 3, useful for detecting, diagnosing or treating tumors,
 PT polycystic kidney disease, diabetic retinopathy, rheumatoid arthritis or
 PT psoriasis.
 XX
 PS Disclosure; Page 138-139; 374pp; English.
 XX
 CC The present invention relates to a novel method for the isolation of
 CC endothelial cells (ECs), and the identification of genes expressed in
 CC normal and tumour ECs. Tumour endothelial marker (TEM), normal
 CC endothelial marker (NEM), and pan-endothelial marker (PEM) genes are
 CC identified in human ECs. The human EC marker proteins and the
 CC polynucleotide sequences encoding them are useful for detecting,
 CC diagnosing or treating tumours as well as polycystic kidney disease,
 CC diabetic retinopathy, rheumatoid arthritis, and psoriasis. They are also
 CC useful for inhibiting neovascularization or tumour angiogenesis, for
 CC inducing an immune response to tumour endothelial cells in a patient, or
 CC for identifying candidate drugs for treating tumours. The present
 CC sequence represents a human TEM or NEM protein of the invention
 XX
 SQ Sequence 564 AA;
 Query Match 100.0%; Score 1526; DB 6; Length 564;
 Best Local Similarity 100.0%; Pred. No. 2.9e-155;
 Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 OGGRRDDGPGACGAGFDLYFLDKSGSVLHNMNEIYFVQOLAKHFTSPQLRMSFIYEST 60
 DB 28 OGGRRDDGPGACGAGFDLYFLDKSGSVLHNMNEIYFVQOLAKHFTSPQLRMSFIYEST 87
 QY 61 RGTILMKLTEDREQIRGLEELQKVLPGDITVMEGGERASEQIYENRQGYRTASVITA 120
 DB 88 RGTILMKLTEDREQIRGLEELQKVLPGDITVMEGGERASEQIYENRQGYRTASVITA 147
 QY 121 LTDEHLEDLFFYSEBRANRSDLGATVYCVGVDFNETOLARLADSDHVPVNDGPOA 180
 DB 148 LTDEHLEDLFFYSEBRANRSDLGATVYCVGVDFNETOLARLADSDHVPVNDGPOA 207
 QY 181 LOGIHSILKSKSCIEILAAEPSTICAGSFQVVRNGNFRARVNDRLVCSFKINDSVTL 240
 DB 208 LOGIHSILKSKSCIEILAAEPSTICAGSFQVVRNGNFRARVNDRLVCSFKINDSVTL 267
 QY 241 NEKPSVEDTYLLCPAPILKEVGKMAALQVSMNDGLSFSSSVITTTTHCSGD 293

DB 268 NEKPSVEDTYLLCPAPILKEVGKMAALQVSMNDGLSFSSSVITTTTHCSGD 320
 RESULT 12
 ADJ70017
 ID ADJ70017 standard; protein; 564 AA.
 XX
 AC ADJ70017;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Human heat mitochondrial protein as a therapeutic target SeqID1823.
 XX
 KM mitochondrial; human; screening assay; diabetes mellitus;
 KM Huntington's disease; osteoarthritis;
 KM Leber's hereditary optic neuropathy; LHON;
 KM mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
 KM myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
 KM neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;
 KM osteopathic; ophthalmological; cytostatic.
 XX
 OS Homo sapiens.
 XX
 PN WC02003087768-A2.
 XX
 PD 23-OCT-2003.
 XX
 PF 04-APR-2003; 2003WO-US010870.
 XX
 PR 12-APR-2002; 2002US-0372843P.
 PR 17-JUN-2002; 2002US-038987P.
 PR 20-SEP-2002; 2002US-0412418P.
 XX
 PA (MITO-) MITOKOR.
 PA (BUCK-) BUCK INST AGE RES.
 XX
 PI Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GW;
 PI Wainrock DE;
 XX
 DR WPI; 2003-845369/78.
 XX
 PT Identifying a mitochondrial target for drug screening assays and for
 PT treating diseases associated with altered mitochondrial function,
 PT comprises detecting a modified polypeptide in a sample and correlating
 PT with the disease.
 XX
 PS Claim 1; SEQ ID NO 1823; 180pp; English.
 XX
 CC This invention relates to novel mitochondrial targets that can be used
 CC for therapeutic intervention in treating a disease associated with
 CC altered mitochondrial function. Specifically, it refers to a method for
 CC identifying proteins of the human heart mitochondrial proteome that are
 CC useful for drug screening assays, as well as therapeutic targets. The
 CC present invention describes a method for identifying such proteins that
 CC can be used in the treatment of various diseases associated with altered
 CC mitochondrial function including diabetes mellitus, Huntington's disease,
 CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
 CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
 CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
 CC compositions have neuroprotective, nootropic, antidiabetic,
 CC anticonvulsant, antirheumatic, osteopathic, ophthalmological and
 CC cytostatic activities. This polypeptide sequence is a human heart
 CC mitochondrial protein of the invention.
 XX
 SQ Sequence 564 AA;
 Query Match 100.0%; Score 1526; DB 7; Length 564;
 Best Local Similarity 100.0%; Pred. No. 2.9e-155;
 Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 OGGRRDDGPGACGAGFDLYFLDKSGSVLHNMNEIYFVQOLAKHFTSPQLRMSFIYEST 60

Db 28 QGRRBDGGPACVGFPLDYFLIDKSGSVLHMNNEYFVEQLAHKFTSPQLRMSFVST 87
 Qy 61 RGTLMKLTEDREQIRQGLEBLQKVLPGDPTVMHEGFERASBOIYYENRQGYRTASVIA 120
 Db 88 RGTLMKLTEDREQIRQGLEBLQKVLPGDPTVMHEGFERASBOIYYENRQGYRTASVIA 147
 Qy 121 LTDEGLHEDLFFYSERANRSDIGAIYVCVGDVFNETOLARIADSKDHYFPVNDGFOA 180
 Db 148 LTDEGLHEDLFFYSERANRSDIGAIYVCVGDVFNETOLARIADSKDHYFPVNDGFOA 207
 Qy 181 LOGIHSILKKSCEIILAAPEPTICAGSFQVYVNGNFRHARNDVRLCSFKINDSVTL 240
 Db 208 LOGIHSILKKSCEIILAAPEPTICAGSFQVYVNGNFRHARNDVRLCSFKINDSVTL 267
 Qy 241 NEKPFVSDTYLLCPAPILKEVGKALQVSNMDGLSFSSSVIITTHCSGD 293
 Db 268 NEKPFVSDTYLLCPAPILKEVGKALQVSNMDGLSFSSSVIITTHCSGD 320
 RESULT 13
 ADR48216 ID ADR48216 standard; protein; 564 AA.
 AC ADR48216;
 XX 18-NOV-2004 (first entry)
 DT 18-NOV-2004 (first entry)
 XX Human tumour endothelial marker 8 precursor protein SEQ:4.
 DE Human tumour endothelial marker 8 precursor protein SEQ:4.
 XX pancreatic cancer-associated transcript; pancreatic cancer; human;
 KW cytostatic; gene therapy; protein therapy;
 KM tumour endothelial marker 8 precursor; TEM8.
 XX Homo sapiens.
 OS Homo sapiens.
 PN WC0004074510-A1.
 XX 02-SEP-2004.
 PD 02-SEP-2004.
 XX 18-FEB-2004; 2004WO-AU000194.
 PF 18-FEB-2004; 2004WO-AU000194.
 XX 18-FEB-2003; 2003AU-00900747.
 PR 18-FEB-2003; 2003AU-00900747.
 XX (GARV-) GARVAN INST MEDICAL RES.
 PA (GARV-) GARVAN INST MEDICAL RES.
 PI Blankin A, Segara D, Henshall S, Sutherland R;
 XX WPI, 2004-635591/61.
 DR N-PSDB; ADR48215.
 XX Detecting pancreatic cancer-associated transcript in a biological sample,
 PT sample for diagnosing or treating the disease, comprises contacting the
 PT sample with a polynucleotide that selectively hybridizes to a specific
 PT sequence.
 XX Claim 70; SEQ ID NO 4; 263bp; English.
 PS The present invention describes a method for detecting a pancreatic
 XX cancer-associated transcript in a biological sample. The method comprises
 CC contacting the biological sample with a polynucleotide that selectively
 CC hybridizes to a sequence at least 80% identical to a sequence as shown in
 CC any one of Tables 3 to 25 in the specification or having the GenBank
 CC Accession Number AF279145. Also described: (1) diagnosing pancreatic
 CC cancer in a human or animal subject being tested, determining the
 CC likelihood that a subject having a pancreatic cancer will survive, or
 CC determining the suitability of a subject having a pancreatic cancer for
 CC surgical resection therapy; (2) detecting a pancreatic cancer-associated
 CC polypeptide in a biological sample; (3) determining the likelihood that a
 CC subject having a pancreatic cancer will survive; and (4) monitoring the
 CC efficacy of a therapeutic treatment of pancreatic cancer. A pancreatic
 CC cancer-associated transcript has cytostatic activity, and can be used in
 CC gene and protein therapy. A pancreatic cancer-associated transcript
 CC polynucleotide, a vector comprising the polynucleotide, an isolated

CC polypeptide or an antibody that binds to the isolated polypeptide can be
 CC used for diagnosing or prognosing pancreatic cancer or for preparing a
 CC treatment for the treatment of pancreatic cancer. The prognostic or
 CC diagnostic methods are useful for the early detection of pancreatic
 CC cancer or its metastases, and for monitoring the progress of disease such
 CC as during remission or following surgery or chemotherapy. The present
 CC sequence represents human tumour endothelial marker 8 precursor (TEM8),
 CC which is used in the exemplification of the present invention.
 XX Sequence 564 AA;
 S0
 Query Match 100.0%; Score 1526; DB 8; Length 564;
 Best Local Similarity 100.0%; Pred. No. 2.9e-155;
 Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 QGRRBDGGPACVGFPLDYFLIDKSGSVLHMNNEYFVEQLAHKFTSPQLRMSFVST 60
 Db 28 QGRRBDGGPACVGFPLDYFLIDKSGSVLHMNNEYFVEQLAHKFTSPQLRMSFVST 87
 Qy 61 RGTLMKLTEDREQIRQGLEBLQKVLPGDPTVMHEGFERASBOIYYENRQGYRTASVIA 120
 Db 88 RGTLMKLTEDREQIRQGLEBLQKVLPGDPTVMHEGFERASBOIYYENRQGYRTASVIA 147
 Qy 121 LTDEGLHEDLFFYSERANRSDIGAIYVCVGDVFNETOLARIADSKDHYFPVNDGFOA 180
 Db 148 LTDEGLHEDLFFYSERANRSDIGAIYVCVGDVFNETOLARIADSKDHYFPVNDGFOA 207
 Qy 181 LOGIHSILKKSCEIILAAPEPTICAGSFQVYVNGNFRHARNDVRLCSFKINDSVTL 240
 Db 208 LOGIHSILKKSCEIILAAPEPTICAGSFQVYVNGNFRHARNDVRLCSFKINDSVTL 267
 Qy 241 NEKPFVSDTYLLCPAPILKEVGKALQVSNMDGLSFSSSVIITTHCSGD 293
 Db 268 NEKPFVSDTYLLCPAPILKEVGKALQVSNMDGLSFSSSVIITTHCSGD 320
 RESULT 14
 AAE01469 ID AAE01469 standard; protein; 403 AA.
 XX AAE01469;
 AC AAE01469;
 XX 17-JUL-2001 (first entry)
 DT 17-JUL-2001 (first entry)
 XX Human gene 4 encoded secreted protein HMLFR02, SEQ ID NO:125.
 DE Human gene 4 encoded secreted protein HMLFR02, SEQ ID NO:125.
 XX Human; secreted protein; proliferative disorder; cancer; tumour;
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
 KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
 KW cardiovascular disorder; angiogenic disorder; kidney disorder;
 KW gastrointestinal disorder; pregnancy-related disorder;
 KW endocrine disorder; infection; wound healing; vulnerability; cell culture;
 KW chemopexis; food additive; gene therapy; binding partner identification;
 KW chromosome 19.
 XX Homo sapiens.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 XX Peptide 1..27
 FT /label= signal_peptide
 FT Protein 28..403
 FT /note= "Mature human secreted protein"
 FT Misc-difference 175
 FT /label= Unknown
 FT /note= "Encoded by GKT"
 FT Misc-difference 320
 FT /label= Unknown
 FT /note= "Encoded by SGT"
 FT Misc-difference 331
 FT /label= Unknown

FT	/note= "Encoded by KTC"
FT	Misc-difference 368
FT	/label= Unknown
FT	/note= "Encoded by WGC"
PN	MO200134626-A1.
PD	17-MAY-2001.
PF	01-NOV-2000; 2000WO-US030045.
PR	05-NOV-1999; 99US-0163581P.
PR	30-JUN-2000; 2000US-0215133P.
PA	(HUMA-) HUMAN GENOME SCI INC.
PI	Ruben SM, Komatroulis GA, Moore PA, Birse CE, Nt J;
DR	WPI, 2001-308778/32.
DR	N-PSDB; AAD05334.
PT	New nucleic acid molecules encoding 28 human secreted proteins for
PT	diagnosing, preventing, treating or ameliorating medical conditions and
PT	used as food additives or preservatives.
PS	Claim 11, Page 505-506; 562pp; English.
XX	AAD05300-AAD05379 represent cDNAs corresponding to 28 human secreted
XX	protein genes, and AAE01436-AAE01513 represent the proteins they encode.
XX	AAE01514-AAE01544 represent human secreted protein fragments or variants.
XX	The genes and their secreted proteins are useful for preventing, treating
XX	or ameliorating medical conditions, e.g., by protein or gene therapy.
XX	Pathological conditions can be diagnosed by determining the amount of the
XX	new protein in a sample or by determining the presence of mutations in
XX	the new genes. Specific uses are described for each of the 28 genes,
XX	based on the tissues in which they are most highly expressed, and include
XX	developing products for the diagnosis or treatment of proliferative
XX	disorders, cancer, tumours, foetal and developmental abnormalities,
XX	haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
XX	diseases (e.g., rheumatoid arthritis), inflammation, allergies,
XX	neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
XX	cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
XX	psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
XX	angiogenic disorders, kidney disorders, gastrointestinal disorders,
XX	pregnancy-related disorders, endocrine disorders, and infections. The
XX	proteins can also be used to aid wound healing and epithelial cell
XX	proliferation, to prevent skin aging due to sunburn, to maintain organs
XX	before transplantation, for supporting cell culture of primary tissues,
XX	to regenerate tissues, to identify their cognate ligands or binding
XX	partners, and in chemotaxis, and can be used as a food additive or
XX	preservative to modify storage properties. Antibodies specific for a
XX	protein of the invention can be used in alleviating symptoms associated
XX	with the disorders mentioned above, and in diagnostic immunoassays e.g.,
XX	radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The
XX	present sequence represents a human secreted protein of the invention
XX	Sequence 403 AA;
XX	Query Match 99.3%; Score 1515; DB 4; Length 403;
XX	Best Local Similarity 99.7%; Pred. No. 2,7e-154;
XX	Matches 291; Conservative 0; Mismatches 1; Indels 0; Gaps 0
QY	1 OGGRREDGPGACGFDLYFIIDKSGSVLHHNMEIYFPYEOIAHKFISPOLMNSPIVST 60
DB	28 OGGRREDGPGACGFDLYFIIDKSGSVLHHNMEIYFPYEOIAHKFISPOLMNSPIVST 87
QY	61 RGTTLMLKLTEDREBQIRQGLEELQKVLPGSDTYMHGFEBSAQIYYENRGYRTASVITA 120
DB	88 RGTTLMLKLTEDREBQIRQGLEELQKVLPGSDTYMHGFEBSAQIYYENRGYRTASVITA 147
QY	121 LTDEGEHLELFFYSEBEARSRDLAIIVCYGKYDNEFNQOLARIADSKHVPVNDGFOA 180
DB	148 LTDEGEHLELFFYSEBEARSRDLAIIVCYGKYDNEFNQOLARIADSKHVPVNDGFOA 207

QY	181	LOGIHSILSKSCIEIIAAEPSTICAGESFOVYVRNGGPRHARVNDVCSFKINDSVTL	240
DB	208	LOGIHSILSKSCIEIIAAEPSTICAGESFOVYVRNGGPRHARVNDVCSFKINDSVTL	267
OY	241	NEKPSVEDTYLLCPAPILKEVGKKALQVSMNDGLSFISSSVITTTTHCSD	292
DB	268	NEKPSVEDTYLLCPAPILKEVGKKALQVSMNDGLSFISSSVITTTTHCSD	319
		RESULT 15	
		ABG63873	
XX	ID	ABG63873 standard; protein; 403 AA.	
XX	AC	ABG63873;	
XX	DT	27-AUG-2002 (first entry)	
XX	DE	Human albumin fusion protein #548.	
KM	XX	Albumin fusion protein; therapeutic protein X; human albumin; HA;	
KM	XX	human serum albumin; HSA; cancer; reproductive disorder;	
KM	XX	digestive disorder; immune disorder; endocrine disorder;	
KM	XX	haematopoietic disorder; neural disorder; connective disorder;	
KM	XX	cytostatic; antifertility; antiinflammatory; antitumor;	
KM	XX	immunomodulator; anti-HIV; antidiabetic; haemostatic; mototropic;	
KM	XX	neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;	
XX	XX	osteopathic; antiarthritic.	
OS	OS	Homo sapiens.	
OS	OS	Synthetic.	
PN	XX	WO200177137-A1.	
PD	XX	18-OCT-2001.	
PF	XX	12-APR-2001; 2001WO-US011988.	
PR	XX	12-APR-2000; 2000US-0229358P.	
PR	XX	25-APR-2000; 2000US-0199384P.	
PR	XX	21-DEC-2000; 2000US-0256931P.	
PA	XX	(HUMA-) HUMAN GENOME SCI INC.	
PI	XX	Rosen CA, Haseltine WA;	
PT	XX	WPI; 2002-010886/01.	
PT	XX	New fusion protein for treating disease e.g. diabetes comprises an	
PS	XX	albumin fused to a therapeutic protein.	
PS	XX	Claim 1; Page 872-873; 2102pp; English.	
CC	XX	The present invention relates to albumin fusion proteins comprising a	
CC	XX	therapeutic protein X and human albumin (HA, also known as human serum	
CC	XX	albumin, HSA). The proteins are useful for treating a disease or disorder	
CC	XX	that may be modulated by therapeutic protein X. The albumin extends the	
CC	XX	shelf-life of protein X, and may increase its biological in vitro/in vivo	
CC	XX	activity. The protein is useful for treating and diagnosing disorders	
CC	XX	such as cancer, reproductive disorders, digestive disorders (e.g. Crohn's	
CC	XX	disease, ulcerative colitis), immune disorders (e.g. acquired	
CC	XX	immunodeficiency syndrome, AIDS), endocrine disorders (e.g. diabetes),	
CC	XX	haematopoietic disorders, neural disorders (e.g. Alzheimer's,	
CC	XX	Parkinson's, Creutzfeldt-Jacob disease, encephalomyelitis, meningitis,	
CC	XX	schizophrenia), and connective disorders (e.g. osteoporosis, arthritis).	
CC	XX	ABG63326-ABG65518 represent albumin fusion proteins of the invention	
XX	XX	Sequence 403 AA;	
Query Match	99.3%	Score 1515; DB 5; Length 403;	
Best Local Similarity	99.7%	Pred. No. 2.7e-154;	
Matches 291; Conservative	0;	Mismatches 1; Indels 0; Gaps 0;	

Qy	1	QGRREDGPA	CYGF	DL	YF	IL	D	K	S	V	L	H	M	N	E	I	Y	F	V	E	O	L	A	H	K	F	I	S	P	O	L	R	M	S	F	I	V	S	T	60										
Db	28	QGRREDGPA	CYGF	DL	YF	IL	D	K	S	V	L	H	M	N	E	I	Y	F	V	E	O	L	A	H	K	F	I	S	P	O	L	R	M	S	F	I	V	S	T	87										
Qy	61	RGTLMKLT	EDRE	QI	RO	G	L	E	I	Q	V	L	P	G	D	T	Y	M	H	E	G	F	E	R	A	S	E	O	I	Y	E	N	R	O	G	Y	R	T	A	S	V	I	A	120						
Db	88	RGTLMKLT	EDRE	QI	RO	G	L	E	I	Q	V	L	P	G	D	T	Y	M	H	E	G	F	E	R	A	S	E	O	I	Y	E	N	R	O	G	Y	R	T	A	S	V	I	A	147						
Qy	121	LTDELHED	LF	F	Y	S	E	R	E	A	N	R	S	R	D	L	G	A	I	V	C	V	K	D	F	N	E	T	O	L	A	R	I	A	D	S	K	D	H	V	F	P	V	N	D	G	F	Q	A	180
Db	148	LTDELHED	LF	F	Y	S	E	R	E	A	N	R	S	R	D	L	G	A	I	V	C	V	K	D	F	N	E	T	O	L	A	R	I	A	D	S	K	D	H	V	F	P	V	N	D	G	F	Q	A	207
Qy	181	LOGIHSIL	KKSC	IE	IL	A	B	P	S	T	I	C	A	G	E	S	F	O	V	V	R	G	N	G	F	R	H	A	R	N	V	D	R	V	L	C	S	F	K	I	N	D	S	V	T	L	240			
Db	208	LOGIHSIL	KKSC	IE	IL	A	B	P	S	T	I	C	A	G	E	S	F	O	V	V	R	G	N	G	F	R	H	A	R	N	V	D	R	V	L	C	S	F	K	I	N	D	S	V	T	L	267			
Qy	241	NEKPSVED	TYL	L	C	P	A	P	I	L	K	E	V	G	M	K	A	L	O	V	S	N	D	G	L	S	F	I	S	S	V	I	I	T	T	H	C	S	D	292										
Db	268	NEKPSVED	TYL	L	C	P	A	P	I	L	K	E	V	G	M	K	A	L	O	V	S	N	D	G	L	S	F	I	S	S	V	I	I	T	T	H	C	S	D	319										

Search completed: June 13, 2005, 19:56:50
 Job time : 112.48 secs

THIS PAGE BLANK (USPTO)

Db 144 PECGQEMDIAFLIDSGSIDOSDFTOMKDFVAKLMGQLASTSTSFSLMOYSNLKTHFT 203
Qy 68 LTEDR-----EQIROGLEBELQKVLPGGDTYMHGEPERASEQIYYENROGYRTA-SVIAL 121
Db 204 FTEFKSSLSPOSVDIAIVQ-----GLTYTASGIQKVVLELPHSKNGARKSAKKILIVI 258
Qy 122 TDGELHEDLFFYSE--REANRSRDGAIVYCVGKD-FNE-TOLARI-----ADSKOHVF 172
Db 259 TDGQKFRDPLEYRHVYIPEAKA---GIRYAIVGDAFREPTALQELNTTIGSAPSQDHVF 315
Qy 173 PVNDGFOLQGIHSLILKSCIEILAAEPSTICAGESFQVNVVNGNFRHARNVD 226
Db 316 KVEN-FVALRSIQROIQEK---IFALGTESSSSSFQHEMSQEGFSALSMD 364

RESULT 2

US-08-485-618-46
; Sequence 46, Application US/08485618
; Patent No. 5728533
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vlieten, Monica
; TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,618
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32797
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1155 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-485-618-46

Query Match 10.9%; Score 166; DB 1; Length 1155;

Best Local Similarity 28.2%; Pred. No. 6e-09; Mismatches 96; Indels 30; Gaps 12;

Matches 66; Conservative 42; Mismatches 96; Indels 30; Gaps 12;

Qy 10 PACYG-GFDLYFLIDSGSV-LHHMNEIYFVBOIAHKFISPOLRMSFIVSTRGTLMK 67
Db 144 PECGQEMDIAFLIDSGSIDOSDFTOMKDFVAKLMGQLASTSTSFSLMOYSNLKTHFT 203

Qy 68 LTEDR-----EQIROGLEBELQKVLPGGDTYMHGEPERASEQIYYENROGYRTA-SVIAL 121
Db 204 FTEFKSSLSPOSVDIAIVQ-----GLTYTASGIQKVVLELPHSKNGARKSAKKILIVI 258
Qy 122 TDGELHEDLFFYSE--REANRSRDGAIVYCVGKD-FNE-TOLARI-----ADSKOHVF 172
Db 259 TDGQKFRDPLEYRHVYIPEAKA---GIRYAIVGDAFREPTALQELNTTIGSAPSQDHVF 315
Qy 173 PVNDGFOLQGIHSLILKSCIEILAAEPSTICAGESFQVNVVNGNFRHARNVD 226
Db 316 KVEN-FVALRSIQROIQEK---IFALGTESSSSSFQHEMSQEGFSALSMD 364

RESULT 3

US-08-362-652-46
; Sequence 46, Application US/08362652
; Patent No. 5766850
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vlieten, Monica
; TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,652
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32391
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1155 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-362-652-46

Query Match 10.9%; Score 166; DB 1; Length 1155;

Best Local Similarity 28.2%; Pred. No. 6e-09; Mismatches 96; Indels 30; Gaps 12;

Matches 66; Conservative 42; Mismatches 96; Indels 30; Gaps 12;

Qy 10 PACYG-GFDLYFLIDSGSV-LHHMNEIYFVBOIAHKFISPOLRMSFIVSTRGTLMK 67
Db 144 PECGQEMDIAFLIDSGSIDOSDFTOMKDFVAKLMGQLASTSTSFSLMOYSNLKTHFT 203
Qy 68 LTEDR-----EQIROGLEBELQKVLPGGDTYMHGEPERASEQIYYENROGYRTA-SVIAL 121
Db 204 FTEFKSSLSPOSVDIAIVQ-----GLTYTASGIQKVVLELPHSKNGARKSAKKILIVI 258
Qy 122 TDGELHEDLFFYSE--REANRSRDGAIVYCVGKD-FNE-TOLARI-----ADSKOHVF 172

Db 259 TDGKFRDPLEYRHVIPAERKA---GIIRYAIVGDAFREPTALQELNTIGSAPSODHV 315
Qy 173 PVNDGFOALOGIHSILKSCIEILAAEPSTICAGESFQVVRGNGFRHARNVD 226
Db 316 KVGN-FVALNLSIQIOIEK---IFAIEGTESRSSSFQHEMSQEGFSSALSM 364

RESULT 4

US-08-605-672-46
Sequence 46, Application US/08605672
Patent No. 5817515
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Seer Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/605,672
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 1155 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-605-672-46

Query Match 10.9%; Score 166; DB 2; Length 1155;

Best Local Similarity 28.2%; Pred. No. 6e-09;

Matches 66; Conservative 42; Mismatches 96; Indels 30; Gaps 12;

Qy 10 PACYG-GFDLYFLIDKSGSV-LHHMNEIYFVEQLAHKFIISPOLRMSFIVSTRTGLTMK 67
Db 144 PECGQEMDIAPFLIDSGSISDQSPFTQMKDFVAKLMQGLASTSTSPSLMOYSLNLTHT 203
Qy 68 LTEDR-----EQIRGIEELQKVLPGSDTYMHEGFERASEQIYYENRQGYRTA-SVIAL 121
Db 204 FTEFKSLSPQSLVDIAIVQ-----GLTYASGIQKVVKELFHSKNGAKSAKILIVI 258
Qy 122 TDGELHEDLFYFSE--REARSRDLGAIIVCVGVXD-FNE-TOLARI-----ADSKOHVF 172
Db 259 TDGKFRDPLEYRHVIPAERKA---GIIRYAIVGDAFREPTALQELNTIGSAPSODHV 315

Qy 173 PVNDGFOALOGIHSILKSCIEILAAEPSTICAGESFQVVRGNGFRHARNVD 226
Db 316 KVGN-FVALNLSIQIOIEK---IFAIEGTESRSSSFQHEMSQEGFSSALSM 364

RESULT 5

US-08-482-293A-46
Sequence 46, Application US/08482293A
Patent No. 5831029
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Seer Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,293A
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 1155 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-482-293A-46

Query Match 10.9%; Score 166; DB 2; Length 1155;

Best Local Similarity 28.2%; Pred. No. 6e-09;

Matches 66; Conservative 42; Mismatches 96; Indels 30; Gaps 12;

Qy 10 PACYG-GFDLYFLIDKSGSV-LHHMNEIYFVEQLAHKFIISPOLRMSFIVSTRTGLTMK 67
Db 144 PECGQEMDIAPFLIDSGSISDQSPFTQMKDFVAKLMQGLASTSTSPSLMOYSLNLTHT 203
Qy 68 LTEDR-----EQIRGIEELQKVLPGSDTYMHEGFERASEQIYYENRQGYRTA-SVIAL 121
Db 204 FTEFKSLSPQSLVDIAIVQ-----GLTYASGIQKVVKELFHSKNGAKSAKILIVI 258
Qy 122 TDGELHEDLFYFSE--REARSRDLGAIIVCVGVXD-FNE-TOLARI-----ADSKOHVF 172
Db 259 TDGKFRDPLEYRHVIPAERKA---GIIRYAIVGDAFREPTALQELNTIGSAPSODHV 315
Qy 173 PVNDGFOALOGIHSILKSCIEILAAEPSTICAGESFQVVRGNGFRHARNVD 226

Db 316 KVGK-FVALRSIQIOEK-----IFAIEGTSSSSSFQHEMSQEGFSSALSMD 364

RESULT 6

US-08-943-363-46
Sequence 46, Application US/08943363
Patent No. 5837478
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
TITLE OF INVENTION: No. 5837478el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Seair Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,363
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 1155 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-943-363-46

Query Match 10.9%; Score 166; DB 2; Length 1155;

Best Local Similarity 28.2%; Pred. No. 6e-09; Matches 66; Conservative 42; Mismatches 96; Indels 30; Gaps 12;

Db 10 PACYG-GFDLYFLDKSGSV-LHMHNEIYFVQOLAHKFIISPOLRMSFIVSTGTTLMK 67
144 PECEQEMDIAPLIDSGSISQSDPTQKDFVAKLMGQLASTSTSPSLMOYSLNLKHTFT 203
68 LTEDR-----EIQRLGELIQKVLPGGDTYMHGEPERASEQIYYENRQGYRTA-SVIAL 121
204 FTEFKSLSPQSLVDIVQLQ-----GLTYASGIQKVVKELFPSKNGARSAKKILIVI 258
122 TDGELHEDLFFYSE--REANRSRDLGAIYVCVQKD-FNE-TQLARI-----ADSKDHF 172
259 TDGQKFRDPLEYRHVYPERAKA---GIRYAIQVGDAREFTALQELNLTIGSAPSODHVF 315
173 PVNDGFQALQGIHSLKSCIEILAAEPSTICAGESFQVYVVGNGFRHARNVD 226
316 KVGK-FVALRSIQIOEK-----IFAIEGTSSSSSFQHEMSQEGFSSALSMD 364

RESULT 7

US-09-193-043-46
Sequence 46, Application US/09193043
Patent No. 6251395
GENERAL INFORMATION:
APPLICANT: Gallatin, Michael W.
TITLE OF INVENTION: No. 6251395el Human 2
FILE REFERENCE: 27866/35004
CURRENT APPLICATION NUMBER: US/09/193,043
CURRENT FILING DATE: 1998-11-16
EARLIER APPLICATION NUMBER: 08/173,497
EARLIER FILING DATE: 1993-12-23
EARLIER APPLICATION NUMBER: 08/286,889
EARLIER FILING DATE: 1994-08-05
EARLIER APPLICATION NUMBER: 08/362,652
EARLIER FILING DATE: 1994-12-21
EARLIER APPLICATION NUMBER: 08/943,363
EARLIER FILING DATE: 1997-10-03
NUMBER OF SEQ ID NOS: 114
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 46
LENGTH: 1155
TYPE: PRT
ORGANISM: Mus musculus
US-09-193-043-46

Query Match 10.9%; Score 166; DB 3; Length 1155;

Best Local Similarity 28.2%; Pred. No. 6e-09; Matches 66; Conservative 42; Mismatches 96; Indels 30; Gaps 12;

Db 10 PACYG-GFDLYFLDKSGSV-LHMHNEIYFVQOLAHKFIISPOLRMSFIVSTGTTLMK 67
144 PECEQEMDIAPLIDSGSISQSDPTQKDFVAKLMGQLASTSTSPSLMOYSLNLKHTFT 203
68 LTEDR-----EIQRLGELIQKVLPGGDTYMHGEPERASEQIYYENRQGYRTA-SVIAL 121
204 FTEFKSLSPQSLVDIVQLQ-----GLTYASGIQKVVKELFPSKNGARSAKKILIVI 258
122 TDGELHEDLFFYSE--REANRSRDLGAIYVCVQKD-FNE-TQLARI-----ADSKDHF 172
259 TDGQKFRDPLEYRHVYPERAKA---GIRYAIQVGDAREFTALQELNLTIGSAPSODHVF 315
173 PVNDGFQALQGIHSLKSCIEILAAEPSTICAGESFQVYVVGNGFRHARNVD 226
316 KVGK-FVALRSIQIOEK-----IFAIEGTSSSSSFQHEMSQEGFSSALSMD 364

NUMBER OF SEQ ID NOS: 114
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO: 46
LENGTH: 1155
TYPE: PRT
ORGANISM: Mus musculus
US-09-688-307A-46

Query Match 10.9%; Score 166; DB 4; Length 1155;
Best Local Similarity 28.2%; Pred. No. 6e-09;
Matches 66; Conservative 42; Mismatches 96; Indels 30; Gaps 12;

QY 10 PACYG-GFDLYFLDKSGSV-LHHMNEIYFVEQLAHKFIISPOLRMSFIVSTRTGLTMK 67
DB 144 PECCGQEMDIAFLIDGSGSIDQSDPTQMKDFVAKLMGQLASTSTSFSLMQYSNLKTHFT 203
QY 68 LTEDR-----EQIRQGLEBLQKVLPGSDTYMHGFEFASQIYYENRQGYRTA-SVIAL 121
DB 204 FTBKSSLSPOSVDALVQLQ-----GLTYTASGIQKVVLELPHSKNGARKSAKKILIVI 258
QY 122 TDGELHEDLPFYSE--REARSRDLGAIYVCVYKD-FNE-TOLARI-----ADSKDHVF 172
DB 259 TDGQKFRDPLEYRHVIVPEAKA---GIIRYAIQVGDAREPTALQELNTIGSAPSQDHVF 315
QY 173 PVNDGFOALQGIHSLKSCIEILAAEPSTICAGESFOVVVRNGFRHARNVD 226
DB 316 KVG-N-FVALRSIQRIQIEK-----IFAIRGTSRSSSSSFQHMMSGEGFSALSMD 364

RESULT 9
US-09-350-259-46
Sequence 46, Application US/09350259
Patent No. 6620915
GENERAL INFORMATION:

APPLICANT: Gallatin, Michael W.
APPLICANT: Van der Vlieten, Monica
TITLE OF INVENTION: No. 6620915el Human 2
FILE REFERENCE: 27866/35004
CURRENT APPLICATION NUMBER: US/09/350,259
CURRENT FILING DATE: 1999-07-08
EARLIER APPLICATION NUMBER: 09/193,043
EARLIER FILING DATE: 1998-11-16
EARLIER APPLICATION NUMBER: 08/173,497
EARLIER FILING DATE: 1993-12-23
EARLIER APPLICATION NUMBER: 08/286,889
EARLIER FILING DATE: 1994-08-05
EARLIER APPLICATION NUMBER: 08/362,652
EARLIER FILING DATE: 1994-12-21
EARLIER APPLICATION NUMBER: 08/943,363
NUMBER OF SEQ ID NOS: 114
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO: 46
LENGTH: 1155
TYPE: PRT
ORGANISM: Mus musculus
US-09-350-259-46

Query Match 10.9%; Score 166; DB 4; Length 1155;
Best Local Similarity 28.2%; Pred. No. 6e-09;
Matches 66; Conservative 42; Mismatches 96; Indels 30; Gaps 12;

QY 10 PACYG-GFDLYFLDKSGSV-LHHMNEIYFVEQLAHKFIISPOLRMSFIVSTRTGLTMK 67
DB 144 PECCGQEMDIAFLIDGSGSIDQSDPTQMKDFVAKLMGQLASTSTSFSLMQYSNLKTHFT 203
QY 68 LTEDR-----EQIRQGLEBLQKVLPGSDTYMHGFEFASQIYYENRQGYRTA-SVIAL 121
DB 204 FTBKSSLSPOSVDALVQLQ-----GLTYTASGIQKVVLELPHSKNGARKSAKKILIVI 258
QY 122 TDGELHEDLPFYSE--REARSRDLGAIYVCVYKD-FNE-TOLARI-----ADSKDHVF 172
DB 259 TDGQKFRDPLEYRHVIVPEAKA---GIIRYAIQVGDAREPTALQELNTIGSAPSQDHVF 315

QY 173 PVNDGFOALQGIHSLKSCIEILAAEPSTICAGESFOVVVRNGFRHARNVD 226
DB 316 KVG-N-FVALRSIQRIQIEK-----IFAIRGTSRSSSSSFQHMMSGEGFSALSMD 364

RESULT 10
US-08-485-618-53

Sequence 53, Application US/08485618
Patent No. 5728533
GENERAL INFORMATION:

APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vlieten, Monica
TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Seear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,618
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32797
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 1161 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-485-618-53

Query Match 10.9%; Score 166; DB 1; Length 1161;
Best Local Similarity 28.2%; Pred. No. 6e-09;
Matches 66; Conservative 42; Mismatches 96; Indels 30; Gaps 12;

QY 10 PACYG-GFDLYFLDKSGSV-LHHMNEIYFVEQLAHKFIISPOLRMSFIVSTRTGLTMK 67
DB 144 PECCGQEMDIAFLIDGSGSIDQSDPTQMKDFVAKLMGQLASTSTSFSLMQYSNLKTHFT 203
QY 68 LTEDR-----EQIRQGLEBLQKVLPGSDTYMHGFEFASQIYYENRQGYRTA-SVIAL 121
DB 204 FTBKSSLSPOSVDALVQLQ-----GLTYTASGIQKVVLELPHSKNGARKSAKKILIVI 258
QY 122 TDGELHEDLPFYSE--REARSRDLGAIYVCVYKD-FNE-TOLARI-----ADSKDHVF 172
DB 259 TDGQKFRDPLEYRHVIVPEAKA---GIIRYAIQVGDAREPTALQELNTIGSAPSQDHVF 315
QY 173 PVNDGFOALQGIHSLKSCIEILAAEPSTICAGESFOVVVRNGFRHARNVD 226

Db 316 KVG-N-FVALRSIORIOEK-----IFAIGTESRSSSSFOHEMSQEGFSSALSMD 364

RESULT 11

US-08-362-652-53
; Sequence 53, Application US/08362652
; Patent No. 5766850
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,652
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32391
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1161 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-362-652-53

Query Match 10.9%; Score 166; DB 1; Length 1161;
Best Local Similarity 28.2%; Pred. No. 6e-09;
Matches 66; Conservative 42; Mismatches 96; Indels 30; Gaps 12;

QY 10 PACYG-GFDLYFLDSSGV-LHHMNEIYFVQLAHKFIISPOLRMSFIVSTRGTTLMK 67
DB 144 PECPQGMIDIAFLIDSGSIDQSDFTQMKDFVAKLMQLASTSTSFSLMQYSNLKTHFT 203
QY 68 LTEDR-----EQIRGLELQKVLPGSDTYMHGFEFASQIYYENRGYRTA-SVIAL 121
DB 204 FTEFKSLSPQSLVDIVQ-----GLTYTASGIQKVVLELPHSKNGARKSAKKILIVI 258
QY 122 TDGELHEDLFFYSE--REANRSRLGAIYVCVQKD-FNE-TOLARI-----ADSKDHVF 172
DB 259 TDQGFRLPDEYRHHVPEAKA---GIRYALGVGDAREPTALQELINTIGSAPSODHVF 315
QY 173 PVNDGFQALQGIHSLIKKSCIEIIAAPSITCAGESFQVYVNGNFRHARNVD 226
DB 316 KVG-N-FVALRSIORIOEK-----IFAIGTESRSSSSFOHEMSQEGFSSALSMD 364

RESULT 12

US-08-605-672-53
; Sequence 53, Application US/08605672
; Patent No. 5817515
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/605,672
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1161 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-605-672-53

Query Match 10.9%; Score 166; DB 2; Length 1161;
Best Local Similarity 28.2%; Pred. No. 6e-09;
Matches 66; Conservative 42; Mismatches 96; Indels 30; Gaps 12;

QY 10 PACYG-GFDLYFLDSSGV-LHHMNEIYFVQLAHKFIISPOLRMSFIVSTRGTTLMK 67
DB 144 PECPQGMIDIAFLIDSGSIDQSDFTQMKDFVAKLMQLASTSTSFSLMQYSNLKTHFT 203
QY 68 LTEDR-----EQIRGLELQKVLPGSDTYMHGFEFASQIYYENRGYRTA-SVIAL 121
DB 204 FTEFKSLSPQSLVDIVQ-----GLTYTASGIQKVVLELPHSKNGARKSAKKILIVI 258
QY 122 TDGELHEDLFFYSE--REANRSRLGAIYVCVQKD-FNE-TOLARI-----ADSKDHVF 172
DB 259 TDQGFRLPDEYRHHVPEAKA---GIRYALGVGDAREPTALQELINTIGSAPSODHVF 315
QY 173 PVNDGFQALQGIHSLIKKSCIEIIAAPSITCAGESFQVYVNGNFRHARNVD 226
DB 316 KVG-N-FVALRSIORIOEK-----IFAIGTESRSSSSFOHEMSQEGFSSALSMD 364

RESULT 13
US-08-482-293A-53
; Sequence 53, Application US/08482293A

```
/ Patent No. 5831029
/ GENERAL INFORMATION:
/ APPLICANT: Gallatin, W. Michael
/ APPLICANT: Van der Vlieten, Monica
/ TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit
/ NUMBER OF SEQUENCES: 103
/ CORRESPONDENCE ADDRESS:
/ ADDRESSER: Marshall, O'Toole, Gerstein, Murray & Borun
/ STREET: 233 South Wacker Drive, 6300 Sear Tower
/ CITY: Chicago
/ STATE: Illinois
/ COUNTRY: United States
/ ZIP: 60606-6402
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/482,293A
/ FILING DATE:
/ CLASSIFICATION: 530
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/173,497
/ FILING DATE: 23-DEC-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/286,889
/ FILING DATE: 5-AUG-1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/362,652
/ FILING DATE: 21-DEC-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Williams Jr., Joseph A.
/ REGISTRATION NUMBER: 38,659
/ REFERENCE/DOCKET NUMBER: 27866/32684
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 312-474-6300
/ TELEFAX: 312-474-0448
/ TELEX: 25-3856
/ INFORMATION FOR SEQ ID NO: 53:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1161 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-482-293A-53

Query Match 10.9%; Score 166; DB 2; Length 1161;
Best Local Similarity 28.2%; Pred. No. 6e-09;
Matches 66; Conservative 42; Mismatches 96; Indels 30; Gaps 12;

QY 10 PACVG-GFDLYFLDKSGSV-LHHMNEIYVFEQLAHKFTSPQARMSFIYSTRGTTLMK 67
DB 144 PEPCGQEMDAFLADSGSIDQSDFTQMKDFVAKMGQLASTSTSFIMQYINILKTHFT 203
QY 68 LTEDR-----EQIRGLELQKVLPGSDTYMHGFEPRASEQIYENRQYRTA-SVIAL 121
DB 204 FTEFKSLSPQSLVDAIVQLQ-----GLTYTASGIQKVVELPHSKNGAKSAKILIVI 258
QY 122 TDGELHDLFFYSE--REARRSDLGAIVYCVGKD-FNE-TOLARI-----ADSKDHF 172
DB 259 TDGQKRPDLPEYHNVIPAEAKA---GIIRYALIGVGAFFRPTALQELINTIGSAPSDHVF 315
QY 173 PVNDGFALOGIHSILKSCIEILAEPTICAGSFQVYVNGNGFRHARNVD 226
DB 316 KVGK-FVALNSIQRIQIEK---IFALGTESSRSSSFQHEMSQEGFSALSMD 364

RESULT 14
US-08-943-363-53
; Sequence 53. Application US/08943363
; Patent No. 5837478
; GENERAL INFORMATION:
; APPLICANT: Van der Vlieten, Monica
```

```
/ APPLICANT: Gallatin, W. Michael
/ APPLICANT: Van der Vlieten, Monica
/ TITLE OF INVENTION: No. 5837478el Human 2 Integrin Alpha Subunit
/ NUMBER OF SEQUENCES: 114
/ CORRESPONDENCE ADDRESS:
/ ADDRESSER: Marshall, O'Toole, Gerstein, Murray & Borun
/ STREET: 233 South Wacker Drive, 6300 Sear Tower
/ CITY: Chicago
/ STATE: Illinois
/ COUNTRY: United States
/ ZIP: 60606-6402
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/943,363
/ FILING DATE:
/ CLASSIFICATION: 530
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/173,497
/ FILING DATE: 23-DEC-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/286,889
/ FILING DATE: 5-AUG-1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/362,652
/ FILING DATE: 21-DEC-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Williams Jr., Joseph A.
/ REGISTRATION NUMBER: 38,659
/ REFERENCE/DOCKET NUMBER: 27866/32684
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 312-474-6300
/ TELEFAX: 312-474-0448
/ TELEX: 25-3856
/ INFORMATION FOR SEQ ID NO: 53:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1161 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-943-363-53

Query Match 10.9%; Score 166; DB 2; Length 1161;
Best Local Similarity 28.2%; Pred. No. 6e-09;
Matches 66; Conservative 42; Mismatches 96; Indels 30; Gaps 12;

QY 10 PACVG-GFDLYFLDKSGSV-LHHMNEIYVFEQLAHKFTSPQARMSFIYSTRGTTLMK 67
DB 144 PEPCGQEMDAFLADSGSIDQSDFTQMKDFVAKMGQLASTSTSFIMQYINILKTHFT 203
QY 68 LTEDR-----EQIRGLELQKVLPGSDTYMHGFEPRASEQIYENRQYRTA-SVIAL 121
DB 204 FTEFKSLSPQSLVDAIVQLQ-----GLTYTASGIQKVVELPHSKNGAKSAKILIVI 258
QY 122 TDGELHDLFFYSE--REARRSDLGAIVYCVGKD-FNE-TOLARI-----ADSKDHF 172
DB 259 TDGQKRPDLPEYHNVIPAEAKA---GIIRYALIGVGAFFRPTALQELINTIGSAPSDHVF 315
QY 173 PVNDGFALOGIHSILKSCIEILAEPTICAGSFQVYVNGNGFRHARNVD 226
DB 316 KVGK-FVALNSIQRIQIEK---IFALGTESSRSSSFQHEMSQEGFSALSMD 364

RESULT 15
US-09-193-043-53
; Sequence 53. Application US/09193043
; Patent No. 6251395
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vlieten, Monica
```

```

1  TITLE OF INVENTION: No. 6251195e1 Human 2
2  FILE REFERENCE: 27866/35004 3
3  CURRENT APPLICATION NUMBER: US/09/193,043 4
4  CURRENT FILING DATE: 1998-11-16 5
5  EARLIER APPLICATION NUMBER: 08/173,497 6
6  EARLIER FILING DATE: 1993-12-23 7
7  EARLIER APPLICATION NUMBER: 08/286,889 8
8  EARLIER FILING DATE: 1994-08-05 9
9  EARLIER APPLICATION NUMBER: 08/362,652 10
10 EARLIER FILING DATE: 1994-12-21 11
11 EARLIER APPLICATION NUMBER: 08/943,363 12
12 EARLIER FILING DATE: 1997-10-03 13
13 NUMBER OF SEQ ID NOS: 114 14
14 SOFTWARE: PatentIn Ver. 2.0 15
15 SEQ ID NO 53 16
16 17
18 19
19 20
20 TYPE: PAT 21
21 ORGANISM: Mus musculus 22
22 OS-09-193-043-53 23

```

Query Match 10.9%; Score 166; DB 3; Length 1161;

```

OY      10 PACYVG--GPDYUFLYLLKSGSV--LHMNNEIYUPEOLAHNEI SPOLRMSYSPVTRGTLTKM 67
Db      144 BECPGQEMDIAFLIDSGSSIDSDPFOKMDPFCVAKLMGLASTSTSPSLMOYSNLTKTHFT 203
OY      68  LTEDR-----EQIRQGLEBLQKVLPGQDTRYVMEIGFERASEQIYYENRQYRTA--SVIAT 121
Db      204 FTFEKSLSLSPQSLVDAIVQLQ-----GLTYTASGIQYKVKELPFSKNGARSAKILIV 258
OY      122 TDGELHEDLEFFYSE--REANRSRDIGALYVCYVKD--FNE--TOLARI-----ADSKDHVF 172
Db      259 TDGQKRPDPLREYRHVPIPEAKA---GIIRYAIQVDAREPPTALQELNTIGSAPSQDHVF 315
OY      173 PVNDGFOALGIIHSLKLKSCIEIILAESTTICAESFOVVVRNGRPHRANND 226
Db      316 KVEN--FVALRSIGROIQEK-----IFAIETBESSRSSSQHEMSEOGESALSMD 364

```

Search completed: June 13, 2005, 20:05:08
Job time : 29.4879 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: June 13, 2005, 20:03:29 ; Search time 93.9469 Seconds
(without alignments)
1195.537 Million cell updates/sec

Title: US-09-970-076-2_COPY_28_320

Perfect score: 1526
Sequence: 1 OGGRRDGGACGCGFDLYF.....DGLSSIVITTHCSGDG 293

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1710399 seqs, 38334425 residues

Total number of hits satisfying chosen parameters: 1710399

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubppaa/US11_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubppaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1526	100.0	403	11 US-09-833-245-621
2	1526	100.0	551	14 US-10-038-307-18
3	1526	100.0	551	14 US-10-201-292-18
4	1526	100.0	564	10 US-09-918-715-187
5	1526	100.0	564	10 US-09-918-715-232
6	1526	100.0	564	14 US-10-301-822-193
7	1526	100.0	564	16 US-10-408-765A-1833
8	1526	100.0	564	16 US-10-474-794-187
9	1526	100.0	564	16 US-10-474-794-232
10	1515	99.3	403	11 US-09-833-245-620
11	1514	99.2	333	10 US-09-796-753-12

12	1514	99.2	333	14 US-10-038-307-2	Sequence 2, App1
13	1514	99.2	333	14 US-10-201-292-2	Sequence 2, App1
14	1514	99.2	345	14 US-10-038-307-24	Sequence 24, App1
15	1514	99.2	345	14 US-10-201-292-24	Sequence 24, App1
16	1514	99.2	564	14 US-10-038-307-20	Sequence 20, App1
17	1514	99.2	564	14 US-10-201-292-20	Sequence 20, App1
18	1509	98.9	562	10 US-09-918-715-194	Sequence 194, App
19	1509	98.9	562	10 US-09-918-715-301	Sequence 301, App
20	1509	98.9	562	16 US-10-474-794-194	Sequence 194, App
21	1509	98.9	562	16 US-10-474-794-301	Sequence 301, App
22	1501	98.4	328	14 US-10-038-307-26	Sequence 26, App1
23	1501	98.4	328	14 US-10-201-292-26	Sequence 26, App1
24	1499.5	98.3	342	14 US-10-038-307-22	Sequence 22, App1
25	1499.5	98.3	342	14 US-10-201-292-22	Sequence 22, App1
26	1493	97.8	543	14 US-10-038-307-14	Sequence 14, App1
27	1493	97.8	543	14 US-10-038-307-16	Sequence 16, App1
28	1493	97.8	543	14 US-10-201-292-14	Sequence 14, App1
29	1493	97.8	543	14 US-10-201-292-16	Sequence 16, App1
30	1483	97.2	543	14 US-10-038-307-10	Sequence 10, App1
31	1483	97.2	543	14 US-10-201-292-10	Sequence 10, App1
32	1478	96.9	534	14 US-10-038-307-12	Sequence 12, App1
33	1478	96.9	534	14 US-10-201-292-12	Sequence 12, App1
34	1413	92.6	529	14 US-10-201-292-36	Sequence 36, App1
35	1299.5	85.2	504	14 US-10-201-292-34	Sequence 34, App1
36	1172	76.8	479	14 US-10-201-292-32	Sequence 32, App1
37	1058	69.3	460	14 US-10-201-292-28	Sequence 28, App1
38	1048	68.7	460	14 US-10-201-292-30	Sequence 30, App1
39	903	59.2	538	13 US-10-047-542-99	Sequence 99, App1
40	786	51.5	488	10 US-09-796-753-52	Sequence 52, App1
41	786	51.5	488	14 US-10-038-307-6	Sequence 6, App1
42	786	51.5	488	14 US-10-201-292-6	Sequence 6, App1
43	786	51.5	488	14 US-10-368-087-16	Sequence 16, App1
44	786	51.5	488	15 US-10-104-047-2639	Sequence 2639, App
45	773.5	50.7	587	9 US-09-764-870-312	Sequence 312, App

ALIGNMENTS

RESULT 1
US-09-833-245-621
Sequence 621, Application US/09833245
Publication No. US20040010134A1
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: PFS46PCT
CURRENT APPLICATION NUMBER: US/09/833,245
CURRENT FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 60/229,358
PRIOR FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: 60/256,931
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/199,384
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 2267
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 621
LENGTH: 403
TYPE: PRT
ORGANISM: Homo sapiens
US-09-833-245-621

Query Match 100.0%; Score 1526; DB 11; Length 403;
Best Local Similarity 100.0%; Pred. No. 1.1e-147;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 OGGRRDGGACGCGFDLYFILDKSGSVLHHNMEIYFVQLAHKFTSPQLRMSFIYFST 60
Db 28 OGGRRDGGACGCGFDLYFILDKSGSVLHHNMEIYFVQLAHKFTSPQLRMSFIYFST 87
Qy 61 RGTTLKLTEDREQIRGLSELOKVLPGSDTYHMEGERASEQIYYENRCQYRTASVITA 120

```

Db      88 RGTTLMLKLTEDREQIRQGLELQKVLPGGDTYMHGFERASEQIYYENRQGYRTASV11A 147
Qy      121 LTDELHEDLFFYSERENRSDLGAIYVCVGDYFNETQIARIADSKDHVFPVNDGFOA 180
Db      148 LTDELHEDLFFYSERENRSDLGAIYVCVGDYFNETQIARIADSKDHVFPVNDGFOA 207
Qy      181 LOGIHSILKKSCIEIIAABPSTTCAGSFQVVRGNGFRRARVNDVRLCSFKINDSVTL 240
Db      208 LOGIHSILKKSCIEIIAABPSTTCAGSFQVVRGNGFRRARVNDVRLCSFKINDSVTL 267
Qy      241 NEKPSVEDTYLLCPAPILKEVGKMAALQVSMNDGLSFSSVITTTTHCSG 293
Db      268 NEKPSVEDTYLLCPAPILKEVGKMAALQVSMNDGLSFSSVITTTTHCSG 320

```

RESULT 2

```

US-10-038-307-18
; Sequence 18, Application US/10038307
; Publication No. US20030134786A1
; GENERAL INFORMATION:
; APPLICANT: James B. ROTTMAN
; APPLICANT: Theresa L. O'KEEFE
; APPLICANT: Engin OZKAYNAK
; APPLICANT: Judith J. HEALEY
; TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
; FILE REFERENCE: 7853-253-999
; CURRENT APPLICATION NUMBER: US/10/038.307
; CURRENT FILING DATE: 2002-06-28
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 551
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-038-307-18

```

```

Query Match      100.0%; Score 1526; DB 14; Length 551;
Best Local Similarity 100.0%; Pred. No. 1.8e-147;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 OGGRRDGGPACYGFDLYFLIDKSGSVLHNMNIYYFVQLAHKFISPOLRMSFIYFST 60
Db      28 OGGRRDGGPACYGFDLYFLIDKSGSVLHNMNIYYFVQLAHKFISPOLRMSFIYFST 87

Qy      61 RGTTLMLKLTEDREQIRQGLELQKVLPGGDTYMHGFERASEQIYYENRQGYRTASV11A 120
Db      88 RGTTLMLKLTEDREQIRQGLELQKVLPGGDTYMHGFERASEQIYYENRQGYRTASV11A 147

Qy      121 LTDELHEDLFFYSERENRSDLGAIYVCVGDYFNETQIARIADSKDHVFPVNDGFOA 180
Db      148 LTDELHEDLFFYSERENRSDLGAIYVCVGDYFNETQIARIADSKDHVFPVNDGFOA 207

Qy      181 LOGIHSILKKSCIEIIAABPSTTCAGSFQVVRGNGFRRARVNDVRLCSFKINDSVTL 240
Db      208 LOGIHSILKKSCIEIIAABPSTTCAGSFQVVRGNGFRRARVNDVRLCSFKINDSVTL 267

Qy      241 NEKPSVEDTYLLCPAPILKEVGKMAALQVSMNDGLSFSSVITTTTHCSG 293
Db      268 NEKPSVEDTYLLCPAPILKEVGKMAALQVSMNDGLSFSSVITTTTHCSG 320

```

RESULT 3

```

US-10-201-292-18
; Sequence 18, Application US/10201292
; Publication No. US20030144193A1
; GENERAL INFORMATION:
; APPLICANT: James B. ROTTMAN
; APPLICANT: Theresa L. O'KEEFE
; APPLICANT: Engin OZKAYNAK
; APPLICANT: Judith J. HEALEY
; TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
; FILE REFERENCE: 7853-253-999
; CURRENT APPLICATION NUMBER: US/10/201.292

```

```

; CURRENT FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 551
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-201-292-18

```

```

Query Match      100.0%; Score 1526; DB 14; Length 551;
Best Local Similarity 100.0%; Pred. No. 1.8e-147;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 OGGRRDGGPACYGFDLYFLIDKSGSVLHNMNIYYFVQLAHKFISPOLRMSFIYFST 60
Db      28 OGGRRDGGPACYGFDLYFLIDKSGSVLHNMNIYYFVQLAHKFISPOLRMSFIYFST 87

Qy      61 RGTTLMLKLTEDREQIRQGLELQKVLPGGDTYMHGFERASEQIYYENRQGYRTASV11A 120
Db      88 RGTTLMLKLTEDREQIRQGLELQKVLPGGDTYMHGFERASEQIYYENRQGYRTASV11A 147

Qy      121 LTDELHEDLFFYSERENRSDLGAIYVCVGDYFNETQIARIADSKDHVFPVNDGFOA 180
Db      148 LTDELHEDLFFYSERENRSDLGAIYVCVGDYFNETQIARIADSKDHVFPVNDGFOA 207

Qy      181 LOGIHSILKKSCIEIIAABPSTTCAGSFQVVRGNGFRRARVNDVRLCSFKINDSVTL 240
Db      208 LOGIHSILKKSCIEIIAABPSTTCAGSFQVVRGNGFRRARVNDVRLCSFKINDSVTL 267

Qy      241 NEKPSVEDTYLLCPAPILKEVGKMAALQVSMNDGLSFSSVITTTTHCSG 293
Db      268 NEKPSVEDTYLLCPAPILKEVGKMAALQVSMNDGLSFSSVITTTTHCSG 320

```

RESULT 4

```

US-09-918-715-187
; Sequence 187, Application US/09918715
; Publication No. US20030017157A1
; GENERAL INFORMATION:
; APPLICANT: Bert Vogelstein
; APPLICANT: Brad St. Croix
; APPLICANT: Kenneth Kinzler
; TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS
; FILE REFERENCE: 1107.00134
; CURRENT APPLICATION NUMBER: US/09/918.715
; CURRENT FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/222,599
; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: 60/224,360
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/282,850
; PRIOR FILING DATE: 2000-04-11
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 187
; LENGTH: 564
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-918-715-187

```

```

Query Match      100.0%; Score 1526; DB 10; Length 564;
Best Local Similarity 100.0%; Pred. No. 1.9e-147;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 OGGRRDGGPACYGFDLYFLIDKSGSVLHNMNIYYFVQLAHKFISPOLRMSFIYFST 60
Db      28 OGGRRDGGPACYGFDLYFLIDKSGSVLHNMNIYYFVQLAHKFISPOLRMSFIYFST 87

Qy      61 RGTTLMLKLTEDREQIRQGLELQKVLPGGDTYMHGFERASEQIYYENRQGYRTASV11A 120
Db      88 RGTTLMLKLTEDREQIRQGLELQKVLPGGDTYMHGFERASEQIYYENRQGYRTASV11A 147

Qy      121 LTDELHEDLFFYSERENRSDLGAIYVCVGDYFNETQIARIADSKDHVFPVNDGFOA 180

```

Db 148 LTDELHEDLFFYSEBRANRSDIGAIVYCVGVDFNETOLARLADSKDHVPVNDGFOA 207
Qy 181 LOGIHSILKKSCEIILAAEPSTICAGSFQVYVNGNGFRHARVNDVLCSEFKINDSVTL 240
Db 208 LOGIHSILKKSCEIILAAEPSTICAGSFQVYVNGNGFRHARVNDVLCSEFKINDSVTL 267
Qy 241 NEKPSVEDTYLLCPAPILKEVGKKAALQVSMNDGLSFISSVIITTHCSDG 293
Db 268 NEKPSVEDTYLLCPAPILKEVGKKAALQVSMNDGLSFISSVIITTHCSDG 320

RESULT 5

US-09-918-715-232
Sequence 232, Application US/09918715
Publication No. US20030017157A1
GENERAL INFORMATION:
APPLICANT: Brad St. Croix
APPLICANT: Bert Vogelstein
APPLICANT: Kenneth Kinzler
TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS
FILE REFERENCE: 1107.00134
CURRENT FILING DATE: 2001-08-01
PRIOR FILING DATE: 2001-08-01
PRIOR APPLICATION NUMBER: 60/222,599
PRIOR FILING DATE: 2000-08-02
PRIOR APPLICATION NUMBER: 60/224,360
PRIOR FILING DATE: 2000-08-11
PRIOR APPLICATION NUMBER: 60/282,850
PRIOR FILING DATE: 2000-04-11
NUMBER OF SEQ ID NOS: 358
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 232
LENGTH: 564
TYPE: PRT
ORGANISM: Homo sapiens
US-09-918-715-232

Query Match 100.0%; Score 1526; DB 10; Length 564;
Best Local Similarity 100.0%; Pred. No. 1.9e-147;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QGGRREDGPGACGCGFDLYFLIDKSGSVLHMMNEIYFVEQLAKHFTSPQLRMSFIVFST 60
Db 28 QGGRREDGPGACGCGFDLYFLIDKSGSVLHMMNEIYFVEQLAKHFTSPQLRMSFIVFST 87
Qy 61 RGTLMKLTEDREIQROGLELQKVLPGDITYMHGEPFRASEOIYENRQGYRTASVILIA 120
Db 88 RGTLMKLTEDREIQROGLELQKVLPGDITYMHGEPFRASEOIYENRQGYRTASVILIA 147
Qy 121 LTDELHEDLFFYSEBRANRSDIGAIVYCVGVDFNETOLARLADSKDHVPVNDGFOA 180
Db 148 LTDELHEDLFFYSEBRANRSDIGAIVYCVGVDFNETOLARLADSKDHVPVNDGFOA 207
Qy 181 LOGIHSILKKSCEIILAAEPSTICAGSFQVYVNGNGFRHARVNDVLCSEFKINDSVTL 240
Db 208 LOGIHSILKKSCEIILAAEPSTICAGSFQVYVNGNGFRHARVNDVLCSEFKINDSVTL 267
Qy 241 NEKPSVEDTYLLCPAPILKEVGKKAALQVSMNDGLSFISSVIITTHCSDG 293
Db 268 NEKPSVEDTYLLCPAPILKEVGKKAALQVSMNDGLSFISSVIITTHCSDG 320

RESULT 6

US-10-301-822-199
Sequence 199, Application US/10301822
Publication No. US20030148410A1
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc.
APPLICANT: Berger, Allison
APPLICANT: Guillemette, Tracy L.
APPLICANT: Kamatkar, Shubhang
APPLICANT: Schlegel, Robert

APPLICANT: Monahan, John E.
APPLICANT: Thibodeau, Stephen N.
APPLICANT: Burgate, Lawrence J.
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
THERAPY OF COLON CANCER
FILE REFERENCE: MEM01-02392RNM
CURRENT FILING DATE: US/10/301,822
PRIOR FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 60/339,971
PRIOR FILING DATE: 2001-12-10
PRIOR APPLICATION NUMBER: US 60/361,978
PRIOR FILING DATE: 2002-03-05
PRIOR APPLICATION NUMBER: US 60/381,988
PRIOR FILING DATE: 2002-05-20
NUMBER OF SEQ ID NOS: 228
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 199
LENGTH: 564
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-301-822-199

Query Match 100.0%; Score 1526; DB 14; Length 564;
Best Local Similarity 100.0%; Pred. No. 1.9e-147;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QGGRREDGPGACGCGFDLYFLIDKSGSVLHMMNEIYFVEQLAKHFTSPQLRMSFIVFST 60
Db 28 QGGRREDGPGACGCGFDLYFLIDKSGSVLHMMNEIYFVEQLAKHFTSPQLRMSFIVFST 87
Qy 61 RGTLMKLTEDREIQROGLELQKVLPGDITYMHGEPFRASEOIYENRQGYRTASVILIA 120
Db 88 RGTLMKLTEDREIQROGLELQKVLPGDITYMHGEPFRASEOIYENRQGYRTASVILIA 147
Qy 121 LTDELHEDLFFYSEBRANRSDIGAIVYCVGVDFNETOLARLADSKDHVPVNDGFOA 180
Db 148 LTDELHEDLFFYSEBRANRSDIGAIVYCVGVDFNETOLARLADSKDHVPVNDGFOA 207
Qy 181 LOGIHSILKKSCEIILAAEPSTICAGSFQVYVNGNGFRHARVNDVLCSEFKINDSVTL 240
Db 208 LOGIHSILKKSCEIILAAEPSTICAGSFQVYVNGNGFRHARVNDVLCSEFKINDSVTL 267
Qy 241 NEKPSVEDTYLLCPAPILKEVGKKAALQVSMNDGLSFISSVIITTHCSDG 293
Db 268 NEKPSVEDTYLLCPAPILKEVGKKAALQVSMNDGLSFISSVIITTHCSDG 320

RESULT 7

US-10-408-765A-1823
Sequence 1823, Application US/10408765A
Publication No. US20040101874A1
GENERAL INFORMATION:
APPLICANT: Ghosh, Soumitra S.
APPLICANT: Fahy, Eoin D.
APPLICANT: Zhang, Bing
APPLICANT: Gibson, Bradford W.
APPLICANT: Taylor, Steven W.
APPLICANT: Glenn, Gary W.
APPLICANT: Warnock, Dale E.
TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
FILE REFERENCE: 660088.465
CURRENT FILING DATE: US/10/408,765A
PRIOR FILING DATE: 2003-04-04
NUMBER OF SEQ ID NOS: 3077
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1823
LENGTH: 564
TYPE: PRT
ORGANISM: Homo sapiens
US-10-408-765A-1823

Query Match 100.0%; Score 1526; DB 16; Length 564;
Best Local Similarity 100.0%; Pred. No. 1.9e-147;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OGGRRDGGPACGCGFDLYFLIDKSGSVLHMNEIYYVEQOLAKHFIISPOLRMSFIYFST 60
DB 28 OGGRRDGGPACGCGFDLYFLIDKSGSVLHMNEIYYVEQOLAKHFIISPOLRMSFIYFST 87
QY 61 RGTLMKLTEDREQIQGLBEIQVLPGGDTYMHGFERASEQIYYENRQGYRTASVYIA 120
DB 88 RGTLMKLTEDREQIQGLBEIQVLPGGDTYMHGFERASEQIYYENRQGYRTASVYIA 147
QY 121 LTDELHEDLFFYSREANRSDIGAIVYCVGVDFNETQARLADSKDHVPVNDGFOA 180
DB 148 LTDELHEDLFFYSREANRSDIGAIVYCVGVDFNETQARLADSKDHVPVNDGFOA 207
QY 181 LOGIHSILKSCIEIILAEPSITCAGSFQVVRGNGFRHARVNDVRLCSFKINDSVTL 240
DB 208 LOGIHSILKSCIEIILAEPSITCAGSFQVVRGNGFRHARVNDVRLCSFKINDSVTL 267
QY 241 NEKPSVEDTYLLCPAPILKEVGKKAALQVSMNDGLSFISSVIITTHCSG 293
DB 268 NEKPSVEDTYLLCPAPILKEVGKKAALQVSMNDGLSFISSVIITTHCSG 320

RESULT 8
US-10-474-794-187
Sequence 187, Application US/10474794
Publication No. US20040213793A1
GENERAL INFORMATION:
APPLICANT: Carson-Walter, Eleanor
APPLICANT: St. Croix, Brad
APPLICANT: Vogelstein, Bert
APPLICANT: Kinzler, Kenneth
TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS
FILE REFERENCE: 1107.00179
CURRENT APPLICATION NUMBER: US/10/474,794
CURRENT FILING DATE: 2003-10-14
PRIOR APPLICATION NUMBER: 60/282,850
PRIOR FILING DATE: 2001-04-11
PRIOR APPLICATION NUMBER: 60/308,829
PRIOR FILING DATE: 2001-08-01
NUMBER OF SEQ ID NOS: 359
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 187
LENGTH: 564
TYPE: PRT
ORGANISM: Homo sapiens
US-10-474-794-187

Query Match 100.0%; Score 1526; DB 16; Length 564;
Best Local Similarity 100.0%; Pred. No. 1.9e-147;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OGGRRDGGPACGCGFDLYFLIDKSGSVLHMNEIYYVEQOLAKHFIISPOLRMSFIYFST 60
DB 28 OGGRRDGGPACGCGFDLYFLIDKSGSVLHMNEIYYVEQOLAKHFIISPOLRMSFIYFST 87
QY 61 RGTLMKLTEDREQIQGLBEIQVLPGGDTYMHGFERASEQIYYENRQGYRTASVYIA 120
DB 88 RGTLMKLTEDREQIQGLBEIQVLPGGDTYMHGFERASEQIYYENRQGYRTASVYIA 147
QY 121 LTDELHEDLFFYSREANRSDIGAIVYCVGVDFNETQARLADSKDHVPVNDGFOA 180
DB 148 LTDELHEDLFFYSREANRSDIGAIVYCVGVDFNETQARLADSKDHVPVNDGFOA 207
QY 181 LOGIHSILKSCIEIILAEPSITCAGSFQVVRGNGFRHARVNDVRLCSFKINDSVTL 240
DB 208 LOGIHSILKSCIEIILAEPSITCAGSFQVVRGNGFRHARVNDVRLCSFKINDSVTL 267
QY 241 NEKPSVEDTYLLCPAPILKEVGKKAALQVSMNDGLSFISSVIITTHCSG 293
DB 268 NEKPSVEDTYLLCPAPILKEVGKKAALQVSMNDGLSFISSVIITTHCSG 320

RESULT 9
US-10-474-794-232
Sequence 232, Application US/10474794
Publication No. US20040213793A1
GENERAL INFORMATION:
APPLICANT: Carson-Walter, Eleanor
APPLICANT: St. Croix, Brad
APPLICANT: Vogelstein, Bert
APPLICANT: Kinzler, Kenneth
TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS
FILE REFERENCE: 1107.00179
CURRENT APPLICATION NUMBER: US/10/474,794
CURRENT FILING DATE: 2003-10-14
PRIOR APPLICATION NUMBER: 60/282,850
PRIOR FILING DATE: 2001-04-11
PRIOR APPLICATION NUMBER: 60/308,829
PRIOR FILING DATE: 2001-08-01
NUMBER OF SEQ ID NOS: 359
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 232
LENGTH: 564
TYPE: PRT
ORGANISM: Homo sapiens
US-10-474-794-232

Query Match 100.0%; Score 1526; DB 16; Length 564;
Best Local Similarity 100.0%; Pred. No. 1.9e-147;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OGGRRDGGPACGCGFDLYFLIDKSGSVLHMNEIYYVEQOLAKHFIISPOLRMSFIYFST 60
DB 28 OGGRRDGGPACGCGFDLYFLIDKSGSVLHMNEIYYVEQOLAKHFIISPOLRMSFIYFST 87
QY 61 RGTLMKLTEDREQIQGLBEIQVLPGGDTYMHGFERASEQIYYENRQGYRTASVYIA 120
DB 88 RGTLMKLTEDREQIQGLBEIQVLPGGDTYMHGFERASEQIYYENRQGYRTASVYIA 147
QY 121 LTDELHEDLFFYSREANRSDIGAIVYCVGVDFNETQARLADSKDHVPVNDGFOA 180
DB 148 LTDELHEDLFFYSREANRSDIGAIVYCVGVDFNETQARLADSKDHVPVNDGFOA 207
QY 181 LOGIHSILKSCIEIILAEPSITCAGSFQVVRGNGFRHARVNDVRLCSFKINDSVTL 240
DB 208 LOGIHSILKSCIEIILAEPSITCAGSFQVVRGNGFRHARVNDVRLCSFKINDSVTL 267
QY 241 NEKPSVEDTYLLCPAPILKEVGKKAALQVSMNDGLSFISSVIITTHCSG 293
DB 268 NEKPSVEDTYLLCPAPILKEVGKKAALQVSMNDGLSFISSVIITTHCSG 320

RESULT 10
US-09-833-245-620
Sequence 620, Application US/09833245
Publication No. US2004010134A1
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: PPS46PCT
CURRENT APPLICATION NUMBER: US/09/833,245
CURRENT FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 60/229,358
PRIOR FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: 60/256,931
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/199,384
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 2267
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 620
LENGTH: 403
TYPE: PRT

```

/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (175)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (320)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (331)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (368)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-833-245-620

```

```

Query Match      99.3%; Score 1515; DB 11; Length 403;
Best Local Similarity 99.7%; Pred. No. 1,5e-146;
Matches 291; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 1 QGGRREDGPGACGCGFDLYFLDKSGSVLHMNEIYFVEQLAHKFTSPQLRMSFIVFST 60
Db 28 QGGRREDGPGACGCGFDLYFLDKSGSVLHMNEIYFVEQLAHKFTSPQLRMSFIVFST 87
Qy 61 RGTLMKLTEDREIQRLQGLLELQVLPFGDPTMHGFERASEQIYYENRQGYRTASVITA 120
Db 88 RGTLMKLTEDREIQRLQGLLELQVLPFGDPTMHGFERASEQIYYENRQGYRTASVITA 147
Qy 121 LTDELHEDLFFYSERANRSDIGALIVYCVGVDPNFETQLARIADSKDHFPPVNDGFOA 180
Db 148 LTDELHEDLFFYSERANRSDIGALIVYCVGVDPNFETQLARIADSKDHFPPVNDGFOA 207
Qy 181 LOGIHSILKSCIEIILAEPSTICAGSFQVVVNGFRRARVDRVLCSEFKINDSVTL 240
Db 208 LOGIHSILKSCIEIILAEPSTICAGSFQVVVNGFRRARVDRVLCSEFKINDSVTL 267
Qy 241 NEKPSVEDTYLLCPAPILKEVGKALQVSMNDGLSFSSSVIITTHCS 292
Db 268 NEKPSVEDTYLLCPAPILKEVGKALQVSMNDGLSFSSSVIITTHCS 319

```

```

RESULT 11
US-09-796-753-12
/ Sequence 12, Application US/09796753
/ Publication No. US2003002798A1
/ GENERAL INFORMATION:
/ APPLICANT: McCarthy, Sean A.
/ TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
/ FILE REFERENCE: 7853-227-999
/ CURRENT APPLICATION NUMBER: US/09/796,753
/ PRIOR FILING DATE: 2001-03-01
/ PRIOR APPLICATION NUMBER: 09/183,175
/ PRIOR FILING DATE: 1998-10-30
/ PRIOR APPLICATION NUMBER: 09/223,094
/ PRIOR FILING DATE: 1998-12-30
/ PRIOR APPLICATION NUMBER: 09/223,546
/ PRIOR FILING DATE: 1998-12-30
/ PRIOR APPLICATION NUMBER: 09/224,246
/ PRIOR FILING DATE: 1998-12-30
/ PRIOR APPLICATION NUMBER: 09/259,388
/ PRIOR FILING DATE: 1999-02-26
/ PRIOR APPLICATION NUMBER: 60/122,458
/ PRIOR FILING DATE: 1999-03-01
/ PRIOR APPLICATION NUMBER: 09/312,359
/ PRIOR FILING DATE: 1999-05-14
/ PRIOR APPLICATION NUMBER: 09/336,536
/ PRIOR FILING DATE: 1999-06-18
/ PRIOR APPLICATION NUMBER: 09/342,687
/ PRIOR FILING DATE: 1999-06-29
/ PRIOR APPLICATION NUMBER: 09/345,464
/ PRIOR FILING DATE: 1999-06-30
/ PRIOR APPLICATION NUMBER: 09/365,164
/ PRIOR FILING DATE: 1999-07-30

```

```

/ PRIOR APPLICATION NUMBER: 09/399,723
/ PRIOR FILING DATE: 1999-09-20
/ PRIOR APPLICATION NUMBER: 09/409,634
/ PRIOR FILING DATE: 1999-09-30
/ PRIOR APPLICATION NUMBER: 09/471,179
/ PRIOR FILING DATE: 1999-12-23
/ PRIOR APPLICATION NUMBER: 09/474,071
/ PRIOR FILING DATE: 1999-12-29
/ PRIOR APPLICATION NUMBER: 09/474,072
/ PRIOR FILING DATE: 1999-12-29
/ PRIOR APPLICATION NUMBER: 09/514,010
/ PRIOR FILING DATE: 2000-02-25
/ PRIOR APPLICATION NUMBER: 09/516,745
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: 09/572,002
/ PRIOR FILING DATE: 2000-05-14
/ PRIOR APPLICATION NUMBER: 09/597,993
/ PRIOR FILING DATE: 2000-06-19
/ PRIOR APPLICATION NUMBER: 09/599,596
/ PRIOR FILING DATE: 2000-06-22
/ PRIOR APPLICATION NUMBER: 09/630,334
/ PRIOR FILING DATE: 2000-07-31
/ PRIOR APPLICATION NUMBER: 09/606,565
/ PRIOR FILING DATE: 2000-06-29
/ PRIOR APPLICATION NUMBER: 09/606,317
/ PRIOR FILING DATE: 2000-06-29
/ PRIOR APPLICATION NUMBER: 09/665,666
/ PRIOR FILING DATE: 2000-09-20
/ PRIOR APPLICATION NUMBER: 09/677,751
/ PRIOR FILING DATE: 2000-09-30
/ NUMBER OF SEQ ID NOS: 162
/ SEQ ID NO 12
/ LENGTH: 333
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-796-753-12

```

```

Query Match      99.2%; Score 1514; DB 10; Length 333;
Best Local Similarity 100.0%; Pred. No. 1,5e-146;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 QGGRREDGPGACGCGFDLYFLDKSGSVLHMNEIYFVEQLAHKFTSPQLRMSFIVFST 60
Db 28 QGGRREDGPGACGCGFDLYFLDKSGSVLHMNEIYFVEQLAHKFTSPQLRMSFIVFST 87
Qy 61 RGTLMKLTEDREIQRLQGLLELQVLPFGDPTMHGFERASEQIYYENRQGYRTASVITA 120
Db 88 RGTLMKLTEDREIQRLQGLLELQVLPFGDPTMHGFERASEQIYYENRQGYRTASVITA 147
Qy 121 LTDELHEDLFFYSERANRSDIGALIVYCVGVDPNFETQLARIADSKDHFPPVNDGFOA 180
Db 148 LTDELHEDLFFYSERANRSDIGALIVYCVGVDPNFETQLARIADSKDHFPPVNDGFOA 207
Qy 181 LOGIHSILKSCIEIILAEPSTICAGSFQVVVNGFRRARVDRVLCSEFKINDSVTL 240
Db 208 LOGIHSILKSCIEIILAEPSTICAGSFQVVVNGFRRARVDRVLCSEFKINDSVTL 267
Qy 241 NEKPSVEDTYLLCPAPILKEVGKALQVSMNDGLSFSSSVIITTHCS 291
Db 268 NEKPSVEDTYLLCPAPILKEVGKALQVSMNDGLSFSSSVIITTHCS 318

```

```

RESULT 12
US-10-018-307-2
/ Sequence 2, Application US/10038307
/ Publication No. US20030134786A1
/ GENERAL INFORMATION:
/ APPLICANT: James B. ROTTMAN
/ APPLICANT: Theresa L. O'KEEFE
/ APPLICANT: Engin OZKANAK
/ APPLICANT: Judith J. HEALEY
/ TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
/ FILE REFERENCE: 7853-253-999

```

;; CURRENT APPLICATION NUMBER: US/10/038,307
;; CURRENT FILING DATE: 2002-06-28
;; NUMBER OF SEQ ID NOS: 26
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 2
;; LENGTH: 333
;; TYPE: PR1
;; ORGANISM: Homo sapiens
US-10-038-307-2

Query Match 99.2%; Score 1514; DB 14; Length 333;
Best Local Similarity 100.0%; Pred. No. 1,5e-146;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QGGRREDGPGACGCGFDLYFLIDKSGSVLHNMNEIYFVEQLAHKFISPOLRMSFIVST 60
Db 28 QGGRREDGPGACGCGFDLYFLIDKSGSVLHNMNEIYFVEQLAHKFISPOLRMSFIVST 87
Qy 61 RGTTLMKLTEDREQIRQGLELQKVLPGGDTYMHGFERASEQIYYENRQGYRTASV11A 120
Db 88 RGTTLMKLTEDREQIRQGLELQKVLPGGDTYMHGFERASEQIYYENRQGYRTASV11A 147
Qy 121 LTDELHEDLFFYSERANRSDLGAIYVCVGDENFETOLAR1ADSXDHFVPVNDGFOA 180
Db 148 LTDELHEDLFFYSERANRSDLGAIYVCVGDENFETOLAR1ADSXDHFVPVNDGFOA 207
Qy 181 LOGIHSILKSCIEILAAEPSTICAGESFOVVVNGNGFRHARVNDVLCSPKINDSVTL 240
Db 208 LOGIHSILKSCIEILAAEPSTICAGESFOVVVNGNGFRHARVNDVLCSPKINDSVTL 267
Qy 241 NEKPSVEDTYLLCPAPILKEVGKKAALQVSMNDGLSFISSVIITTHCS 291
Db 268 NEKPSVEDTYLLCPAPILKEVGKKAALQVSMNDGLSFISSVIITTHCS 318

RESULT 13
US-10-201-292-2
;; Sequence 2, Application US/10201292
;; Publication No. US20030144193A1
;; GENERAL INFORMATION:
;; APPLICANT: James B. ROTTMAN
;; APPLICANT: Theresa L. O'KEEFE
;; APPLICANT: Engin OZKAYNAK
;; APPLICANT: Judith J. HEALEY
;; TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
;; FILE REFERENCE: 7853-253-999
;; CURRENT APPLICATION NUMBER: US/10/201,292
;; CURRENT FILING DATE: 2003-02-14
;; NUMBER OF SEQ ID NOS: 36
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 2
;; LENGTH: 333
;; TYPE: PR1
;; ORGANISM: Homo sapiens
US-10-201-292-2

Query Match 99.2%; Score 1514; DB 14; Length 333;
Best Local Similarity 100.0%; Pred. No. 1,5e-146;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QGGRREDGPGACGCGFDLYFLIDKSGSVLHNMNEIYFVEQLAHKFISPOLRMSFIVST 60
Db 28 QGGRREDGPGACGCGFDLYFLIDKSGSVLHNMNEIYFVEQLAHKFISPOLRMSFIVST 87
Qy 61 RGTTLMKLTEDREQIRQGLELQKVLPGGDTYMHGFERASEQIYYENRQGYRTASV11A 120
Db 88 RGTTLMKLTEDREQIRQGLELQKVLPGGDTYMHGFERASEQIYYENRQGYRTASV11A 147
Qy 121 LTDELHEDLFFYSERANRSDLGAIYVCVGDENFETOLAR1ADSXDHFVPVNDGFOA 180
Db 148 LTDELHEDLFFYSERANRSDLGAIYVCVGDENFETOLAR1ADSXDHFVPVNDGFOA 207
Qy 181 LOGIHSILKSCIEILAAEPSTICAGESFOVVVNGNGFRHARVNDVLCSPKINDSVTL 240

Db 208 LOGIHSILKSCIEILAAEPSTICAGESFOVVVNGNGFRHARVNDVLCSPKINDSVTL 267
Qy 241 NEKPSVEDTYLLCPAPILKEVGKKAALQVSMNDGLSFISSVIITTHCS 291
Db 268 NEKPSVEDTYLLCPAPILKEVGKKAALQVSMNDGLSFISSVIITTHCS 318

RESULT 14
US-10-038-307-24
;; Sequence 24, Application US/10038307
;; Publication No. US20030134786A1
;; GENERAL INFORMATION:
;; APPLICANT: James B. ROTTMAN
;; APPLICANT: Theresa L. O'KEEFE
;; APPLICANT: Engin OZKAYNAK
;; APPLICANT: Judith J. HEALEY
;; TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
;; FILE REFERENCE: 7853-253-999
;; CURRENT APPLICATION NUMBER: US/10/038,307
;; CURRENT FILING DATE: 2002-06-28
;; NUMBER OF SEQ ID NOS: 26
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 24
;; LENGTH: 345
;; TYPE: PR1
;; ORGANISM: Homo sapiens
US-10-038-307-24

Query Match 99.2%; Score 1514; DB 14; Length 345;
Best Local Similarity 100.0%; Pred. No. 1,6e-146;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QGGRREDGPGACGCGFDLYFLIDKSGSVLHNMNEIYFVEQLAHKFISPOLRMSFIVST 60
Db 28 QGGRREDGPGACGCGFDLYFLIDKSGSVLHNMNEIYFVEQLAHKFISPOLRMSFIVST 87
Qy 61 RGTTLMKLTEDREQIRQGLELQKVLPGGDTYMHGFERASEQIYYENRQGYRTASV11A 120
Db 88 RGTTLMKLTEDREQIRQGLELQKVLPGGDTYMHGFERASEQIYYENRQGYRTASV11A 147
Qy 121 LTDELHEDLFFYSERANRSDLGAIYVCVGDENFETOLAR1ADSXDHFVPVNDGFOA 180
Db 148 LTDELHEDLFFYSERANRSDLGAIYVCVGDENFETOLAR1ADSXDHFVPVNDGFOA 207
Qy 181 LOGIHSILKSCIEILAAEPSTICAGESFOVVVNGNGFRHARVNDVLCSPKINDSVTL 240
Db 208 LOGIHSILKSCIEILAAEPSTICAGESFOVVVNGNGFRHARVNDVLCSPKINDSVTL 267
Qy 241 NEKPSVEDTYLLCPAPILKEVGKKAALQVSMNDGLSFISSVIITTHCS 291
Db 268 NEKPSVEDTYLLCPAPILKEVGKKAALQVSMNDGLSFISSVIITTHCS 318

RESULT 15
US-10-201-292-24
;; Sequence 24, Application US/10201292
;; Publication No. US20030144193A1
;; GENERAL INFORMATION:
;; APPLICANT: James B. ROTTMAN
;; APPLICANT: Theresa L. O'KEEFE
;; APPLICANT: Engin OZKAYNAK
;; APPLICANT: Judith J. HEALEY
;; TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
;; FILE REFERENCE: 7853-253-999
;; CURRENT APPLICATION NUMBER: US/10/201,292
;; CURRENT FILING DATE: 2003-02-14
;; NUMBER OF SEQ ID NOS: 36
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 24
;; LENGTH: 345
;; TYPE: PR1
;; ORGANISM: Homo sapiens

US-10-201-292-24

Query Match 99.2%; Score 1514; DB 14; Length 345;
Best Local Similarity 100.0%; Pred. No. 1.6e-146;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	OGGRREDGGPACVCGFDLPFLIDKSGSVLHHMNEIYYFVEQLAHKFI	SPQLRMSFIVEST	60
DB	28	OGGRREDGGPACVCGFDLPFLIDKSGSVLHHMNEIYYFVEQLAHKFI	SPQLRMSFIVEST	87
QY	61	RGTTLMLKLTEDREQIROGLBELQKVLPGGDTYMHGFEFRASEQI	YYENRQGYRTASVILIA	120
DB	88	RGTTLMLKLTEDREQIROGLBELQKVLPGGDTYMHGFEFRASEQI	YYENRQGYRTASVILIA	147
QY	121	LTDGELHEDLFFYGEREARNRSRDIGAIYVCVKDFNETQLARI	ADSKDHVPVNDGFQA	180
DB	148	LTDGELHEDLFFYGEREARNRSRDIGAIYVCVKDFNETQLARI	ADSKDHVPVNDGFQA	207
QY	181	LOGIHSILKKSCEIILAEPSTICAGESFQVYVRGNGFRHARVNDVYL	CSFKINDSVTL	240
DB	208	LOGIHSILKKSCEIILAEPSTICAGESFQVYVRGNGFRHARVNDVYL	CSFKINDSVTL	267
QY	241	NEKPFVEDTYLCPAPILKEVGKKAALQVSMNDGLSFISSVIITTHCS		291
DB	268	NEKPFVEDTYLCPAPILKEVGKKAALQVSMNDGLSFISSVIITTHCS		318

Search completed: June 13, 2005, 20:36:50
Job time : 94.9469 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 19:47:08 ; Search time 21.0213 Seconds

(without alignments)

1341.095 Million cell updates/sec

Title: us-09-970-076-2_copy_28_320

Sequence: 1 QGGRREDGPGACYGGFDFY.....DGLSFSSSVITTTCSGDG 293

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR79:.*
2: PIR2:.*
3: PIR3:.*
4: PIR4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	145.5	9.5	1153	1 RWHUB	cell surface glyco
2	143.5	9.4	1163	1 RWHUB	cell surface glyco
3	143	9.4	3124	2 A40020	collagen alpha 1(X
4	141	9.2	1153	2 S00551	leukocyte surface
5	139	9.1	724	2 A48569	antigen Em100 - Bi
6	137	9.0	1747	2 A45974	collagen alpha 1(X
7	134.5	8.8	712	2 A45638	immunodominant mic
8	131	8.6	760	1 C2MS	classical-compleme
9	130	8.5	1857	2 S31212	collagen alpha 1(X
10	130	8.5	1888	2 S78476	collagen alpha 1(X
11	128	8.4	3051	2 A42373	hypothetical prote
12	118	7.7	929	2 I51027	type XII collagen
13	113.5	7.4	496	2 A37979	cartilage matrix p
14	113.5	7.4	764	1 BBHU	complement factor
15	111.5	7.3	2944	2 A54849	collagen alpha 1(V
16	110.5	7.2	761	1 BBMS	complement factor
17	109.5	7.2	1029	1 S21369	collagen alpha 2(V
18	108.5	7.1	500	2 S66522	cartilage matrix p
19	107.5	7.0	493	2 A33809	cartilage matrix p
20	105.5	6.9	567	2 T28797	hypothetical prote
21	103	6.7	747	2 I51579	complement factor
22	100.5	6.6	574	2 A46283	sporocyste surface
23	100	6.6	460	2 T23087	hypothetical prote
24	99.5	6.5	597	2 S33578	top protein - fru1
25	99	6.5	559	2 S04531	thrombospondin-rel
26	99	6.5	698	2 B96958	dnax protein (heat
27	99	6.5	817	2 T50240	kinesin-like prote
28	99	6.5	917	2 S09646	collagen alpha 2(V
29	99	6.5	1018	1 CGH02A	collagen alpha 2(V

30	98.5	6.5	537	2 T04822	hypothetical prote
31	98	6.4	741	2 T46488	hypothetical prote
32	98	6.4	3176	2 CGH03A	collagen alpha 3(V
33	97.5	6.4	642	2 H81185	dnax protein NMB05
34	97	6.4	3137	2 A37797	collagen alpha 3(V
35	96.5	6.3	676	2 T47637	hypothetical prote
36	96	6.3	272	2 A55348	integrin alpha-1 -
37	96	6.3	689	2 F84811	probable retrovira
38	95.5	6.3	334	2 AF1166	transcription regu
39	95.5	6.3	932	2 JCS953	inter-alpha-inhibi
40	95.5	6.3	1179	2 A53213	integrin alpha-E c
41	94.5	6.2	1151	2 A45226	integrin alpha-1 c
42	94	6.2	843	2 A40970	undulin 1 - human
43	94	6.1	843	2 A40970	hypothetical prote
44	92.5	6.1	642	2 B81917	probable chaperone
45	92	6.0	292	2 B83736	transposase (04) B

ALIGNMENTS

RESULT 1

cell surface glycoprotein CD11b precursor [validated] - human

N/Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein Ma
eukocyte integrin alpha chain; neutrophil adherence receptor alpha chain

C/Species: Homo sapiens (man)

C/Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 09-Jul-2004

C/Accession: A31108; A26915; A41600; A30892; A32218; A46526; A26091; I52567

R/Corbi. A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.

J. Biol. Chem. 263, 12403-12411, 1988

A/Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD

B.

A/Reference number: A31108; MUID:88315033; PMID:2457584

A/Accession: A31108

A/Molecule type: mRNA

A/Residues: 1-1153 <COR>

A/Cross-references: UNIPROT:P11215; GB:J03925; NID:G187284; PIDN:AAA59544.1; PID:G307148

A/Note: part of this sequence was confirmed by protein sequencing

R/Arnaut, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.

J. Cell Biol. 106, 2153-2158, 1988

A/Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor Mac

A/Reference number: A28915; MUID:88257215; PMID:2454931

A/Accession: A28915

A/Molecule type: mRNA

A/Residues: 1-499, 501-965, 'P', 967-1153 <ARN>

A/Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:G186935; PIDN:AAA5944.1

A/Note: the authors translated the codon TAC for residue 1129 as Thr

A/Note: part of this sequence, including the amino end of the mature protein, was confir

R/Shelley, C.S.; Arnaut, M.A.

Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991

A/Title: The promoter of the CD11b gene directs myeloid-specific and developmentally regu

A/Reference number: A41600; MUID:92073318; PMID:1683702

A/Accession: A41600

A/Molecule type: DNA

A/Residues: 1-9 <SHE>

A/Cross-references: GB:M76724; NID:G180018; PIDN:AAA54410.1; PID:G553215

R/Arnaut, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.

A/Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhesion

A/Reference number: A94193; MUID:88190151; PMID:2833753

A/Accession: A30892

A/Molecule type: mRNA

A/Residues: 917-1042 <AR2>

A/Cross-references: GB:M18044

R/Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.

Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989

A/Title: cDNA sequence for the alpha subunit of the human neutrophil adherence receptor

A/Reference number: A32218; MUID:89098893; PMID:2563162

A/Accession: A32218

A/Molecule type: mRNA

A/Residues: 9-1153 <HIC>

A/Cross-references: GB:J04145; NID:G189068; PIDN:AAA59903.1; PID:G386975

A>Note: part of the protein was confirmed by protein sequencing
R.Fleming, U.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.
J. Immunol. 150, 480-490, 1993

A>Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-in-
n during evolution.

A:Reference number: A46526; MUID:93123748; PMID:8419480

A>Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-499,501-1153 <FE>

A:Cross-references: GB:S52227; NID:G263047; PIDN:AAB24621.1; PID:G263049

A>Note: the last three bases of intron 13, CAG, are included in some but not all mature

A:Note: sequence extracted from NCBI backbone (NCBIF:121963)

R.Pierce, M.W.; Remold-O'Donnell, E.; Todd III, R.F.; Arnaout, M.A.
Biochim. Biophys. Acta 874, 368-371, 1986

A>Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across sp-

A:Reference number: A90664; MUID:87076671; PMID:3539202

A:Accession: A26091

A:Molecule type: protein

A:Residues: 17-31 <PI>

A:Experimental source: granulocytes

R.Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G.
Blood 79, 865-870, 1992

A>Title: Characterization of the myeloid-specific CD11b promoter.

A:Reference number: 152567; MUID:92144986; PMID:1346576

A:Accession: 152567

A>Status: translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-9 <RES>

A:Cross-references: GB:M64477; NID:G180184; PIDN:AA51960.1; PID:G553219

C:Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1

C:Genetics:

A:Gene: GDB:ITGM; CR3A

A:Cross-references: GDB:120599; OMIM:120980

A:Map position: 16p11.2-16p11.2

A:Note: Promoter containe a GATA motif and two Spl consensus binding sites

C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homoc-

C:Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; mag-

F:1-16/Domai: signal sequence #status predicted <SIG>

F:17-1153/Product: cell surface glycoprotein CD11b #status experimental <MAT>

F:17-1108/Domai: extracellular #status predicted <EXT>

F:148-318/Domai: von Willebrand factor type A repeat homology <WMA>

F:465-473/Region: calcium/magnesium binding #status predicted

F:530-538/Region: calcium/magnesium binding #status predicted

F:593-601/Region: calcium/magnesium binding #status predicted

F:1109-1134/Domai: transmembrane #status predicted <TM>

F:1155-1153/Domai: intracellular #status predicted <INT>

F:186,240,391,469,693,697,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Binding

```

N:Alternate names: leukocyte adhesion receptor p150,95 alpha chain
C:Species: Homo sapiens (man)
C>Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text_change 09-Jul-2004
C:Accession: A36584; A35543; S00864
R:Cordbl, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 12750-12751, 1990
A:Reference number: A36584
A:Contents: erratum
A:Accession: A36584
A:Molecule type: DNA
A:Residues: 1-1163 <COR>
A:Cross-references: UNIPROT:P20702
A>Note: This revision to the sequence from reference A35543 includes the carboxyl end
R:Cordbl, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 2782-2788, 1990
A>Title: Genomic structure of an integrin alpha subunit, the leukocyte p150,95 molecule.
A:Reference number: A35543; NUID:90153906; PMID:2303426
A:Molecule type: DNA
A:Accession: A35543
A:Residues: 1-834 <CO2>
A>Note: This sequence has been revised in reference A36584
R:Cordbl, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.
EMBO J. 6, 4023-4028, 1987
A>Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocyte
A:Reference number: S00864; NUID:88166645; PMID:3327667
A:Accession: S00864
A:Molecule type: mRNA
A:Residues: 1-755, 'L', 757-1163 <CO3>
A:Cross-references: GB:M81695; EMBL:Y00093; NID:g487829; PIDN:AAA59180.1; PID:g487830
A>Note: part of this sequence was confirmed by protein sequencing
C:Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on myeloid cells
C:Genetics:
A:Gene: GDB:ITGAX; CD11C
A:Cross-references: GDB:119759; OMIM:151510
A:Map position: 16p11.2-16p11.2
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homo-
C:Keywords: calcium, cell adhesion; glycoprotein; heterodimer; magnesium; tandem repeat;
F1-19/Dominant: signal sequence #status predicted <SIG>
F20-1163/Product: cell surface glycoprotein CD11c #status predicted <MAT>
F20-1107/Dominant: extracellular #status predicted <EXT>
F149-319/Dominant: von Willebrand factor type A repeat homology <VWA4>
F1108-1133/Dominant: transmembrane #status predicted <TM>
F1134-1163/Dominant: intracellular #status predicted <INT>
F611,89,392,697,735,899,939,1050/Binding site: carbohydrate (Asn) (covalent) #status pre-
dicted
Query Match 9.4%; Score 143.5; DB 1; Length 1163;
Best Local Similarity 24.8%; Pred. No. 0.0024;
Matches 55; Conservative 43; Mismatches 89; Indels 35; Gaps 10;

QY 17 DLYFLDKSGSV-LHHMNEIYYVEQLAHKFIISPOLMGSTIVSTRGTTLMLKTEDR-- 73
|:::||||| :::: |::: |||||
DB 151 DIVFLIDGGSSISRNFATMFNVRAVISQFORPSDFSLMQFSNKQTFHFEFRRTS 210
|:::||||| :::: |::: |||||
QY 74 ---QRQGLEBIOKVLPGDVTYMHGEFBRASEDIYYENRCGYRAS-VIALTQGLIED 129
|:::||||| :::: |::: |||||
DB 211 NFLSLIASVHQD-----GETYATATAIQNVAHRLPHASYGARPDATKILLYITDSKEGD 265
|:::||||| :::: |::: |||||
QY 130 LEFYSEREANRRDLGAIVYCVG-----KDNFTOLARIAD-SKHVPFPVNDG 177
|:::||||| :::: |::: |||||
DB 266 SLDDYD-VIPMDAGGIIRYAIGVGLAFOGRNSWKELND-----IAKPSEGHTFKYVD- 318
|:::||||| :::: |::: |||||
QY 178 FOALQGIHSILKKSCTEIIAAEPSTTCAGESQOVVVRNGGF 219
|:::||||| :::: |::: |||||
DB 319 FPLAKDIOQMUTKEK---IFAIEGTETTSSTSSFELEMAOEGF 356
|:::||||| :::: |::: |||||

RESULT 3
A40020 collagen alpha 1(XII) chain precursor - chicken
N:Alternate names: fibrochmerin
C:Species: Gallus gallus (chicken)
C>Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 09-Jul-2004
C:Accession: A40020; A34485; B34485; A28037;S23814; S22254; S28811
```

R.Yamagata, M.; Yamada, K.M.; Yamada, S.S.; Shimomura, T.; Tanaka, H.; Nishida, Y.; Obay
 U. Cell Biol. 115, 209-221, 1991
 A>Title: The complete primary structure of type XII collagen shows a chimeric molecule w
 nous region of type IX collagen, and short collagenous domains with an Arg-Gly-Asp site.
 A/Reference number: A40020, MUID:92011862; PMID:1918137
 A/Accession: A40020
 A/Molecule type: mRNA
 A/Residues: 1-3124 <YAM>
 A/Cross-references: UNIPROT:P13944; GB:D00824; NID:9222810; PIDN:BA00701.1; PID:9222811
 A/Note: In the authors' translation residues 1216-1219 are shown after residue 1235 and,
 R.Gordon, M.K.; Gerecke, D.R.; Dublet, B.; van der Rest, M.; Olsen, B.R.
 J. Biol. Chem. 264, 19772-19778, 1989
 A>Title: Type XII collagen. A large multidomain molecule with partial homology to type I
 A/Reference number: A34485; MUID:90062079; PMID:2584192
 A/Accession: A34485
 A/Molecule type: mRNA
 A/Residues: 2456-2758, 'A', 2760-2802, 'F', 2804-2976, 'F', 2978-3124 <GOR>
 A/Cross-references: EMBL:J05137; NID:9211284; PIDN:AAA48635.1; PID:9211285
 A/Accession: B34485
 A/Molecule type: protein
 A/Residues: 2772-2792;2846-2873 <GOR2>
 R.Gordon, M.K.; Gerecke, D.R.; Olsen, B.R.
 Proc. Natl. Acad. Sci. U.S.A. 86, 6040-6044, 1987
 A>Title: Type XII collagen: distinct extracellular matrix component discovered by cDNA c
 A/Reference number: A28037; MUID:8731590; PMID:3476925
 A/Accession: A28037
 A/Molecule type: mRNA
 A/Residues: 2960-2976, 'P', 2978-3074, 'AG' <GOR3>
 A/Cross-references: EMBL:M17375; NID:9211649; PIDN:AAA48718.1; PID:9211650
 A/Note: this sequence has been revised in reference A34485
 R.Koch, M.; Bernasconi, C.; Chiquet, M.
 Eur. J. Biochem. 207, 847-856, 1992
 A>Title: A major oligomeric fibroblast proteoglycan identified as a novel large form of
 A/Reference number: S23814; MUID:92362621; PMID:1323460
 A/Accession: S23814
 A/Molecule type: protein
 A/Residues: 'X', 1333, 'Q', 1335-1347;1394-1928;2504, 'X', 2506, 'X', 2508-2511, 'X', 2513-2517 <
 R.Dublet, B.; van der Rest, M.
 J. Biol. Chem. 267, 17724-17727, 1992
 A>Title: Type XII collagen is expressed in embryonic chick tendons. Isolation of pepsin-
 A/Reference number: S22254; MUID:88087065; PMID:3121603
 A/Accession: S22254
 A/Molecule type: protein
 A/Residues: 2831-2832, 'T', 2834, 'R', 2836-2843;3002-3014 <DUB>
 R.Tueb, J.; Tueb, B.
 Biochim. Biophys. Acta 1171, 97-98, 1992
 A>Title: The two splice variants of collagen XII share a common 5' end.
 A/Reference number: S28811; MUID:93042014; PMID:1420368
 A/Accession: S28811
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-24,1189-1257, 'S', 1259-1263, 'E', 1265-1280 <TRU>
 A/Cross-references: EMBL:X67327
 C/Genetics:
 A/Intons: 2845/3; 2863/3; 2887/3; 2899/3; 2922/1; 2985/1; 3008/1; 3065/1
 C/Keywords: alternative splicing; cell binding; coll binding; coll; connective tissue; disulfid
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-3124/Product: collagen alpha 1(XII) chain #status predicted <MAT>
 F:24-1189-1257/Product: collagen alpha 1(XII) chain short splice form #status predicted
 F:24-114/Domain: I11A #status predicted <I11A>
 F:24-105/Domain: fibronectin type III repeat homology <FN3A>
 F:137-301/Domain: von Willebrand factor type A repeat homology <VWA1>
 F:133-425/Domain: I11B #status predicted <I11B>
 F:133-414/Domain: fibronectin type III repeat homology <FN3B>
 F:437-601/Domain: von Willebrand factor type A repeat homology <VWA2>
 F:629-1178/Domain: I11C #status predicted <I11C>
 F:630-711/Domain: fibronectin type III repeat homology <FN3C>
 F:721-802/Domain: fibronectin type III repeat homology <FN3D>
 F:812-895/Domain: fibronectin type III repeat homology <FN3E>
 F:905-986/Domain: fibronectin type III repeat homology <FN3F>
 F:999-1076/Domain: fibronectin type III repeat homology <FN3G>
 F:1086-1169/Domain: fibronectin type III repeat homology <FN3H>
 F:1197-1361/Domain: von Willebrand factor type A repeat homology <VWA3>

F:1384-2295/Domain: I11D #status predicted <I11D>
 F:1384-1465/Domain: fibronectin type III repeat homology <FN3I>
 F:1474-1557/Domain: fibronectin type III repeat homology <FN3J>
 F:1566-1647/Domain: fibronectin type III repeat homology <FN3K>
 F:1655-1738/Domain: fibronectin type III repeat homology <FN3L>
 F:1756-1838/Domain: fibronectin type III repeat homology <FN3M>
 F:1847-1928/Domain: fibronectin type III repeat homology <FN3N>
 F:1937-2019/Domain: fibronectin type III repeat homology <FN3O>
 F:2028-2110/Domain: fibronectin type III repeat homology <FN3P>
 F:2119-2199/Domain: fibronectin type III repeat homology <FN3Q>
 F:2207-2299/Domain: fibronectin type III repeat homology <FN3R>
 F:2325-2490/Domain: von Willebrand factor type A repeat homology <VWA4>
 F:2438-2440/Region: cell adhesion #status predicted
 F:2509-2750/Domain: IXP, homologous to NC4 domain of type IX collagen #status predicted
 F:2751-2902/Domain: collagenous COL2 #status predicted <COL2>
 F:2893-2901/Region: cell attachment (R-G-D) motif
 F:2903-2945/Domain: non-collagenous NC2 #status predicted <NC2>
 F:2946-3048/Domain: collagenous COL1 #status predicted <COL1>
 F:3049-3124/Domain: non-collagenous NC1 #status predicted <NC1>
 F:32,1006,1032,1044,1512,1767,2210,2273,2532,2683/Binding site: carboxylate (Asn) (cova)
 F:2780,2789,2836,2842,2860,2869,3004,3007/Modified site: hydroxyproline (Pro) #statu

Query Match 9.4%; Score 143; DB 2; Length 3124;

Best Local Similarity 26.1%; Pred. No. 0.0097; Mismatches 97; Indels 36; Gaps 14;

Matches 65; Conservative 51; Mismatches 97; Indels 36; Gaps 14;

17 DLYFLDKSGSV-LHNNNEIYFVEOLAHKF-ISP-OLRMSFVSTRTTLMKLTEDRE 73

Db 439 DVEFLVGVGSISIGANVVKRAFLFVYKSFELSPKVOISLVQYSR--DPHHEFSLNRY 496

Qy 74 QIRGSELEOKLP--GGDTYMEGFERASEQYIENRQGR--TASVITLTDGEIHEH 129

Db 497 NRVYDIIQAIINTFPRGSGSTNTGKAMTYVEKVFVTSK--GSRPVVPRVMTLITDGK--SSD 554

Qy 130 LFFPSESEARNSDGLAIVYCVGKDPNFRQIARID--SKDHFVPNDGFQALQGIHS 187

Db 555 AF--KBPATKLRADYEVFAVGVKQAVRTELEINAPPAETHYTED--FDAFQIRISFE 610

Qy 188 ILKSCIEIILAEPRSTICAGESFOVVRNGGFRARV--DRVLGCFKINDSVTLNEKP 244

Db 611 LTQGVCLRI-----EQELAIKRSYPAKAMVSDVTSDFKSKWASGE-- 657

Qy 245 FSVEDTYLL 253

Db 658 ---EKSYLI 663

RESULT 4

S00551 leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse

N:Alternate names: complement-3 receptor alpha chain

C:Species: Mus musculus (house mouse)

C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004

C/Accession: S00551; I59078

R:PyTela, R.

EMBO J. 7, 1371-1378, 1988

A>Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the ir

A/Reference number: S00551; MUID:88312584; PMID:3044779

A/Accession: S00551

A/Molecule type: DNA

A/Residues: 1-1153 <PYT>

A/Cross-references: UNIPROT:P05555; EMBL:X07640; NID:952982; PIDN:CAA30479.1; PID:952983

A/Note: the authors translated the codon CAC for residue 569 as Gln

R:Sastre, L.; Roman, J.M.; Teglow, D.B.; Dreyer, W.J.; Gee, C.E.; Larson, R.S.; Roberts,

Proc. Natl. Acad. Sci. U.S.A. 83, 5644-5648, 1986

A>Title: A partial genomic DNA clone for the alpha subunit of the mouse complement recept

A/Reference number: I59078; MUID:86287312; PMID:2942940

A/Accession: I59078

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 11-44 <RES>

A/Cross-references: GB:M14293; NID:9189993; PIDN:AAA39484.1; PID:9554193

C/Genetics:

A:Gene: Mac-1
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
C:Keywords: cell adhesion; glycoprotein; transmembrane protein
F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-1153/Product: leukocyte surface glycoprotein Mac-1 alpha chain #status experimental
F:146-318/Domain: von Willebrand factor type A repeat homology <VWA2>
F:1106-1129/Domain: transmembrane #status predicted <TM>

Query Match 9.2%; Score 141; DB 2; Length 1153;

Best Local Similarity 24.2%; Pred. No. 0.0038;
Matches 71; Conservative 50; Mismatches 96; Indels 76; Gaps 16;

Qy 17 DLYFLDKSGSVLHHNNEIYFVEQLAHKFI--PQLRMSFIVSTRTGTLTKLTED-- 71
Db 150 DIVFLIDSGSI-----NNIDF---QIKMKEFVSTWEOQPKSKTLFS-----LMQYSDER 197
Qy 72 ----REQRQGLHELOKVL--CGDTYMHGFERASEQIYVE--NRCGYTASVIALT 122
Db 198 IHFTFNDKRNKPSRSHSPITKQLNGRTKTASGIRKVRRELPHKNGARENAKILVIT 257
Qy 123 DGELEHDLFFYSE--REANRSRDLGAIYVCVGD--FNETOLARIAD-----SKDHVP 173
Db 258 DGEKRGDPLDKVDYREADRA---GVIRYVGVGNAPFKPSRRRLDTIAPKAGENVFQ 314
Qy 174 VNDGFQALOGIITHSLKSCIEILAAPEPTICAGESFQVVRNGFRRHARNDVRLCSFK 233
Db 315 V-DNFEALNTIIONLOEK-----IFAIEGTGTGTSSTSFHEMSQEGF----- 355
Qy 234 INDSVTLNKXPFVSVDYTLCLCAPILKEVG---MKALQVSMNDGLSFISS 282
Db 356 -SASITSN-----GPIUGSVGFPMAGAFLYTYSKDKVTFINTT 393

RESULT 5

A:Antigen Em100 - Elmeria maxima

C:Species: Elmeria maxima

C>Date: 01-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

C:Accession: A48569

R:Parasitology, L.; Hug, D.; Humbelin, M.; Weber, G.

MOL. Biochem. Parasitol. 57, 171-174, 1993

A:Title: Sequence of a major Elmeria maxima antigen homologous to the Elmeria tenella m

A:Reference number: A48569; PMID:93149203; PMID:8426611

A:Accession: A48569

A:Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-724 <PAS>

A:CROSS-References: UNIPROT:004588; GB:M99058; NID:9158890; PID:9158891

A:Note: Sequence extracted from NCBI Backbone (NCBI:123776, NCBI:P:123777)

F:45-218/Domain: von Willebrand factor type A repeat homology <VWA2>

F:236-317/Domain: fibronectin type III repeat homology <FN3A>

F:132-432/Domain: fibronectin type III repeat homology <FN3B>

F:433-493/Domain: fibronectin type III repeat homology <FN3C>

F:494-556/Domain: fibronectin type III repeat homology <FN3D>

F:560-610/Domain: fibronectin type I repeat homology <FN3E>

Query Match 9.1%; Score 139; DB 2; Length 724;

Best Local Similarity 26.6%; Pred. No. 0.003; Indels 34; Gaps 14;

Matches 59; Conservative 40; Mismatches 89; Indels 34; Gaps 14;

Qy 12 CYGFDPLVFLDKSGSV-LHHNNEIYFVEQLAHKFI--PQLRMSFIVSTRTGTLTKLT 68
Db 42 CTRLLDVMVLVDESSIGTSIGSVYKVRISFNPAQMPISPDVDRVGLVTFGSAVTRMDL 101
Qy 69 TEDRQIQGLFELQKVL--CGDTYMHGFERASEQIYENRQGYR--TASVIALTDG 124
Db 102 SDSRQNDMLAAAKKPYAAGSTYTHGLAKA--EELTFSPQGRNAPMPLVMTDG 160
Qy 125 ELHEDLFFYSERE-----ANRSRDLGAIYVCVGV--KOFNETOLARIA--DSKHV--FP- 173
Db 161 A-----SSRSQTLAAEKLRNRGVIIIVLVGTVGNASACRSIAGCTSDTVECPR 212

Qy 174 -VNDGFQALOGIITHSLKSCIEI---LAAERSTI--CAGE 208
Db 213 YLSNMGVSSQINGIITKACDLADAVCSWSEYGPCEGE 254

RESULT 6

A:Antigen alpha 1(XIV) chain precursor, short form 2 - chicken

C:Species: Gallus gallus (chicken)

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C:Accession: A45974; S30085; S22916; S17035; S20833

F:Gene: D.R.; Foley, J.W.; Castagnola, P.; Gemari, M.; Dublet, B.; Cancedda, R.; Lin

J. Biol. Chem. 268, 12177-12184, 1993

A:Title: Type XIV collagen is encoded by alternative transcripts with distinct 5' regions

A:Reference number: A45974; PMID:93280195; PMID:8505337

A:Accession: A45974

A:Status: preliminary

A:Molecule type: mRNA; protein

A:Residues: 1-1747 <GER>

A:CROSS-References: UNIPROT:P32018

A:Note: sequence inconsistent with the nucleotide translation

A:Reference number: S22916; PMID:92339443; PMID:1339349

A:Accession: S22916

A:Molecule type: mRNA

A:Residues: 1472-1660 <APT>

A:CROSS-References: EMBL:X65122; NID:962871; PID:CAA46238.1; PID:9938175

R:Trueb, J.; Trueb, B.

Eur. J. Biochem. 207, 549-557, 1992

A:Title: Type XIV collagen is a variant of undulin.

A:Reference number: S22916; PMID:92339443; PMID:1339349

A:Accession: S22916

A:Molecule type: mRNA

A:Status: preliminary

A:Residues: 286-494, 'Q', 496-834, 'A', 836-1119, 'KL', 1122-1402, 1409-1439 <TRU>

R:Gordon, M.K.; Castagnola, P.; Dublet, B.; Linsemayr, T.F.; van der Rest, M.; Mayne, F

Eur. J. Biochem. 201, 333-338, 1991

A:Title: Cloning of a cDNA for a new member of the class of fibril-associated collagens

A:Reference number: S17035; PMID:92037585; PMID:1935930

A:Accession: S17035

A:Molecule type: mRNA

A:Residues: 1472-1659 <GOR1>

A:Accession: S20833

A:Molecule type: protein

A:Keywords: alternative splicing; coiled coil; extracellular matrix; glycoprotein; trimer

F:40-204/Domain: von Willebrand factor type A repeat homology <VWA1>

F:236-317/Domain: fibronectin type III repeat homology <FN3A>

F:132-432/Domain: fibronectin type III repeat homology <FN3B>

F:433-493/Domain: fibronectin type III repeat homology <FN3C>

F:494-556/Domain: fibronectin type III repeat homology <FN3D>

F:560-610/Domain: fibronectin type I repeat homology <FN3E>

F:716-798/Domain: fibronectin type III repeat homology <FN3F>

F:924-1089/Domain: fibronectin type III repeat homology <FN3G>

F:1111-1352/Domain: non-collagenous NC4 #status predicted <NC4>

F:1511-1553/Domain: non-collagenous NC2 #status predicted <NC2>

F:1554-1659/Domain: triple helical domain COL1 #status predicted <COL1>

Query Match 9.0%; Score 137; DB 2; Length 1747;

Best Local Similarity 25.4%; Pred. No. 0.014; Indels 26; Gaps 12;

Matches 63; Conservative 45; Mismatches 114; Indels 26; Gaps 12;

Qy 17 DLYFLDKSGSV-LHHNNEIYFVEQL--AHKFI--PQLRMSFIVSTRTGTLTKLT--E 70
Db 926 DIVFLVDSGWSIGDNNKILISFLYSTVGALDKIGPGTQVALIIFQEDDPRTEKLANAYK 985
Qy 71 DREQIQGLFELQKVLPGDTYMHGFERASEQIYENRQGYR--TASVIALTDGELHE 128

```

Db      986 TKEITLLEIQQI--AYKGGNTTKGAIKHAKE-VLFTGEGARKKIPKVLVTTIDGRSQD 1042
Qy      129 DLFFYSEREANRSRL-GAIVCVGVKDFNETQARIAD--SKDHVFPVNDGFQALQGI 185
Db      1043 DV-----NKVREMGDLDFSPFALIVADADYSELVNISKSESRVFFVDD-FDAFTIE 1096
Qy      186 HSLKSCIEILAEAPSTTCAGESQVAVVVRGNGFPHANNVDRVLCSPKINDSVTLNEKP 245
Db      1097 DELTIFVETASATCPLVFKQDNFA-----GFKMEMFGILVEKFSADGVSMEPGT 1150
Qy      246 SVEDTYLL 253
Db      1151 NVYPCIRL 1158

```

RESULT 7

```

A45638
Immunodominant microneme protein Epi100 - Eimeria tenella
C/Species: Eimeria tenella
C/Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C/Accession: A45638
R/Tomley, F.M.; Clarke, L.E.; Kawazoe, U.; Djikema, R.; Kok, J.J.
Mol. Biochem. Parasitol. 49, 277-288, 1991
A/Title: Sequence of the gene encoding an immunodominant microneme protein of Eimeria tenella
A/Reference number: A45638, MUID:92131064, PMID:1775171
A/Accession: A45638
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-712 <TOM>
A/Cross-reference: UNIPROT:O43981; GB:A032905; GB:M73495; NID:92707732; PIDN:AAD03350.
A/Note: Sequence extracted from NCBI backbone (NCBI:77752, NCBI:77756)
F/148-216/Domain: von Willebrand factor type A repeat homology <VWA1>
F/238-286/Domain: thrombospondin type 1 repeat homology <THR1>
F/309-371/Domain: thrombospondin type 1 repeat homology <THR2>
F/372-432/Domain: thrombospondin type 1 repeat homology <THR3>
F/433-493/Domain: thrombospondin type 1 repeat homology <THR4>
F/494-556/Domain: thrombospondin type 1 repeat homology <THR5>
F/560-610/Domain: thrombospondin type 1 repeat homology <THR6>

```

```

Query Match      8.8%; Score 134.5; DB 2; Length 712;
Best Local Similarity 24.6%; Pred. No. 0.0068;
Matches 55; Conservative 40; Mismatches 88; Indels 41; Gaps 11;

Qy      12 CYGSGFDLYFLDKSGSV-LHHMNEIYFVEQLAHKF-ISPQ-LRMSFVSTRTIMKL 68
Db      45 CTSLLDWLVVDESGISCTSNFRKVRQRIEDPNVSMPTSPEDVAVGLTPT----- 96
Qy      69 TEDREQIRQGLEELQKVLPG-----GDTVMHSGFERASEQIYYENQGYR-- 113
Db      97 ---RSKYRWMLSDPKATPSPALISAASLSYSTGVTYTHYGLQDA-KKLVDYVNAARNN 152
Qy      114 TASYIIALTOGELHEDLFFYSEREANRSRDIGAIYVCVGV-KDENEQVLARIADSKOHVF 172
Db      153 VPKVLVLTWTDAA--SNLPSQTRSSAAMLRDGAIVVVLGVGSGVSSSCRSIACSTSN 210
Qy      173 P-----VNDGFQALQGIHSLKSCIEILAEAPSTTC-CAE 208
Db      211 PRYQSNWSNVTQVNGIIRAKCDLAKDAVCEMSSEYGPVGE 254

```

RESULT 8

```

C2MS
classical-complement-pathway C3/C5 convertase (BC 3.4.21.43) C2 component precursor - mc
N/Alternate names: C3 convertase; C5 convertase; complement C2
C/Species: Mus musculus (house mouse)
C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C/Accession: A38876, B36593, I54429
R/Ishikawa, N.; Nonaka, M.; Wetzel, R.A.; Colten, H.R.
Submitted to GenBank, January 1991
A/Reference number: A38876
A/Accession: A38876
A/Molecule type: DNA

```

```

A/Residues: 1-760 <IS2>
A/Cross-reference: UNIPROT:P21180; GB:M57891; GB:J05661; NID:9192436; PIDN:AAA63294.1;
R/Ishikawa, N.; Nonaka, M.; Wetzel, R.A.; Colten, H.R.
J. Biol. Chem. 265, 19040-19046, 1990
A/Title: Murine complement C2 and factor B genomic and cDNA cloning reveals different mechanisms
A/Reference number: A36593, MUID:91035430, PMID:2229060
A/Accession: B36593
A/Molecule type: mRNA
A/Residues: 1-760 <ISH>
A/Cross-reference: EMBL:M57891; NID:9192436; PIDN:AAA63294.1; PID:9192437
R/Falut, A.; Wakeland, R.K.; McConnell, T.J.; Gitlin, J.; Whitehead, A.S.; Colten, H.R.
Immunogenetics 25, 290-298, 1987
A/Title: DNA polymorphism of MHC III genes in inbred and wild mouse strains.
A/Reference number: I54429, MUID:87192938, PMID:2883115
A/Accession: I54429
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 660-677, R', 679, 681-723, G', 725 <RBS>
A/Cross-reference: GB:M16271, NID:9199289; PIDN:AAA39562.1; PID:9199290
A/Genetics:
A/Intons: 16/1; 91/1; 153/1; 212/1; 245/3; 290/3; 337/1; 384/1; 414/1; 461/1; 492/3; 53/3
A/Complex: The proenzyme forms a complex with C4a and is activated by cleavage into C2a,
C/Function:
A/Description: cleaves complement C3 and complement C5 alpha chains
A/Pathway: complement classical pathway
C/Superfamily: complement B/C2; complement factor H repeat homology; trypsin homology;
C/Keywords: alternative splicing; complement classical pathway; duplication; glycoprotein;
F/1-18/Domain: signal sequence #status predicted <SIG>
F/19-250/Domain: complement C2b fragment #status predicted <C2b>
F/251-260/Domain: complement factor H repeat homology <FH1>
F/261-270/Domain: complement factor H repeat homology <FH2>
F/271-280/Domain: complement factor H repeat homology <FH3>
F/281-290/Domain: complement factor H repeat homology <FH4>
F/291-300/Domain: complement factor H repeat homology <FH5>
F/301-310/Domain: complement factor H repeat homology <FH6>
F/311-320/Domain: complement factor H repeat homology <FH7>
F/321-330/Domain: complement factor H repeat homology <FH8>
F/331-340/Domain: complement factor H repeat homology <FH9>
F/341-350/Domain: complement factor H repeat homology <FH10>
F/351-360/Domain: complement factor H repeat homology <FH11>
F/361-370/Domain: complement factor H repeat homology <FH12>
F/371-380/Domain: complement factor H repeat homology <FH13>
F/381-390/Domain: complement factor H repeat homology <FH14>
F/391-400/Domain: complement factor H repeat homology <FH15>
F/401-410/Domain: complement factor H repeat homology <FH16>
F/411-420/Domain: complement factor H repeat homology <FH17>
F/421-430/Domain: complement factor H repeat homology <FH18>
F/431-440/Domain: complement factor H repeat homology <FH19>
F/441-450/Domain: complement factor H repeat homology <FH20>
F/451-460/Domain: complement factor H repeat homology <FH21>
F/461-470/Domain: complement factor H repeat homology <FH22>
F/471-480/Domain: complement factor H repeat homology <FH23>
F/481-490/Domain: complement factor H repeat homology <FH24>
F/491-500/Domain: complement factor H repeat homology <FH25>
F/501-510/Domain: complement factor H repeat homology <FH26>
F/511-520/Domain: complement factor H repeat homology <FH27>
F/521-530/Domain: complement factor H repeat homology <FH28>
F/531-540/Domain: complement factor H repeat homology <FH29>
F/541-550/Domain: complement factor H repeat homology <FH30>
F/551-560/Domain: complement factor H repeat homology <FH31>
F/561-570/Domain: complement factor H repeat homology <FH32>
F/571-580/Domain: complement factor H repeat homology <FH33>
F/581-590/Domain: complement factor H repeat homology <FH34>
F/591-600/Domain: complement factor H repeat homology <FH35>
F/601-610/Domain: complement factor H repeat homology <FH36>
F/611-620/Domain: complement factor H repeat homology <FH37>
F/621-630/Domain: complement factor H repeat homology <FH38>
F/631-640/Domain: complement factor H repeat homology <FH39>
F/641-650/Domain: complement factor H repeat homology <FH40>
F/651-660/Domain: complement factor H repeat homology <FH41>
F/661-670/Domain: complement factor H repeat homology <FH42>
F/671-680/Domain: complement factor H repeat homology <FH43>
F/681-690/Domain: complement factor H repeat homology <FH44>
F/691-700/Domain: complement factor H repeat homology <FH45>
F/701-710/Domain: complement factor H repeat homology <FH46>
F/711-720/Domain: complement factor H repeat homology <FH47>
F/721-730/Domain: complement factor H repeat homology <FH48>
F/731-740/Domain: complement factor H repeat homology <FH49>
F/741-750/Domain: complement factor H repeat homology <FH50>
F/751-760/Domain: complement factor H repeat homology <FH51>
F/761-770/Domain: complement factor H repeat homology <FH52>
F/771-780/Domain: complement factor H repeat homology <FH53>
F/781-790/Domain: complement factor H repeat homology <FH54>
F/791-800/Domain: complement factor H repeat homology <FH55>
F/801-810/Domain: complement factor H repeat homology <FH56>
F/811-820/Domain: complement factor H repeat homology <FH57>
F/821-830/Domain: complement factor H repeat homology <FH58>
F/831-840/Domain: complement factor H repeat homology <FH59>
F/841-850/Domain: complement factor H repeat homology <FH60>
F/851-860/Domain: complement factor H repeat homology <FH61>
F/861-870/Domain: complement factor H repeat homology <FH62>
F/871-880/Domain: complement factor H repeat homology <FH63>
F/881-890/Domain: complement factor H repeat homology <FH64>
F/891-900/Domain: complement factor H repeat homology <FH65>
F/901-910/Domain: complement factor H repeat homology <FH66>
F/911-920/Domain: complement factor H repeat homology <FH67>
F/921-930/Domain: complement factor H repeat homology <FH68>
F/931-940/Domain: complement factor H repeat homology <FH69>
F/941-950/Domain: complement factor H repeat homology <FH70>
F/951-960/Domain: complement factor H repeat homology <FH71>
F/961-970/Domain: complement factor H repeat homology <FH72>
F/971-980/Domain: complement factor H repeat homology <FH73>
F/981-990/Domain: complement factor H repeat homology <FH74>
F/991-1000/Domain: complement factor H repeat homology <FH75>

```

```

Query Match      8.6%; Score 131; DB 1; Length 760;
Best Local Similarity 23.2%; Pred. No. 0.014;
Matches 74; Conservative 56; Mismatches 115; Indels 74; Gaps 16;

Qy      14 GGFVLYILDKSSGVLLHNMIEY-YFVEQLAHKFIQSLR--NSFYESTRTIMKLTE 70
Db      258 GHLNLYLLDLASQSVTEKDFDFKSAELMVERIFSEFVAVTVAITFASOPTIMSLIS 317
Qy      71 DREQIRQGLEELQKVLPGDPTVMHSGFERASEQIYYE-----NROGYRTAS- 116
Db      318 ERSG-----DVEVITSLSASAYKHENATGANTYEVLLIRVYSMMQTQMDRLGMETSAN 371
Qy      117 ----VIALTDELHEDLFFYSEREANRSRDIGAI-----YVCVGV-----KD 155
Db      372 KEIRHTIILITDK--SNMGDSPKAVTRIRELSTIEQNRDYLIDYIAIVGKLDVDMKE 429
Qy      156 FNETQARIADSDKHDPVVDGFOALQGIHSLKSCIEILAEAPSTTCAGSPQVYVR 215
Db      430 LNE--LGSKDGSRHHAITLDA--KALQQLFEHNLVDK-----TDTICG-----V 472
Qy      216 GNGFRHANNVDRV--LCSFKINDSVTLNEKPSVEEDTYLLCPAPILKEVGMK--AAQVLS 271
Db      473 GNSANASDERFTPWQVTFKPKSKETCGS--LISQWVTLAAHCFHDIMQEDHMLKRVN 530
Qy      272 MND-----GLSPISSVYI 285
Db      531 VQDPTSGHKEFLVEDYII 549

```

RESULT 9

```

S31212
collagen alpha 1(XIV) chain precursor, short form - chicken
C/Species: Gallus gallus (chicken)

```

C>Date: 11-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 15-Sep-2003
C/Accession: S31212
R/Maechli, C.; Trueb, U.; Kessler, B.; Winterhalter, K.H.; Trueb, B.
Eur. J. Biochem. 212, 483-490, 1993
A/Title: Complete primary structure of chicken collagen XIV.
A/Reference number: S31211; MUID:93185668; PMID:8444186
A/Accession: S31212
A/Structure: nucleic acid sequence not shown; translation not shown
A/Molecule type: mRNA
A/Residues: 1-1857 <NAME>
A/Cross-references: EMBL:X70792; NID:G288874; PIDN:CA50064.1; PID:G288875
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
C/Genetics:
A/Gene: Coll14A1
C/Keywords: alternative splicing; coiled coil; extracellular matrix; glycoprotein; trimer
F/1-28/Domain: signal sequence #status predicted <SIG>
F/29-1857/Product: collagen alpha 1(XIV) chain, short form #status predicted <MAT>
F/29-110/Domain: fibronectin type III repeat homology <FN3A>
F/156-330/Domain: von Willebrand factor type A repeat homology <VWA1>
F/352-433/Domain: fibronectin type III repeat homology <FN3B>
F/442-525/Domain: fibronectin type III repeat homology <FN3C>
F/534-614/Domain: fibronectin type III repeat homology <FN3D>
F/623-707/Domain: fibronectin type III repeat homology <FN3E>
F/741-823/Domain: fibronectin type III repeat homology <FN3F>
F/832-914/Domain: fibronectin type III repeat homology <FN3G>
F/922-1009/Domain: fibronectin type III repeat homology <FN3H>
F/1040-1205/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 8.5%; Score 130; DB 2; Length 1857;
Best Local Similarity 25.0%; Pred. No. 0.055; Indels 26; Gaps 12;
Matches 62; Conservative 44; Mismatches 116;
Qy 17 DLVFLDKSGSV-LHHNNEIYFVEQL-AHKFISPO-LRMSFIVSTRGTLMKLT--E 70
Db 1042 DLVFLVDSGWSIGDDNFNKIISFLYSTVGALDKIGDGTQVAILQFSDDPRTFKLNAYK 1101
Qy 71 DREOIRQGLEBLOKVLPGDVTYMEGFERASEQIYYENRQGYRTA--SVIALLDGELEH 128
Db 1102 TKETLLBAIQOI--AVKGGNTKTGKAIKHARE-VLFTGEAGMKRGIPKVLVITDGRSQD 1158
Qy 129 DLFFYSERANRNDL-GAIVYCVGVKDPNFTQALRIAD--SKDHYFPVNDGQALQGI 185
Db 1159 DV-----NKVSRMQLDGSPFAIGVADADYSELVNISSKPSRRHVPFVD--FDAFTKIE 1212
Qy 186 HSILKSCIEILAEPTICAGESFQVNVVRNGFRHARNDVRLCSFKINDSVTLNEKPF 245
Db 1213 DELITFVETASATCPVLVFKDGDKLA-----GFMGMMPGLVKEKFSALDGVSMERGTF 1266
Qy 246 SVEDTYLL 253
Db 1267 NVYPCYRL 1274

RESULT 10
S78476
collagen alpha 1(XIV) chain precursor, long form - chicken
C/Species: Gallus gallus (chicken)
C/Date: 11-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 09-Jul-2004
C/Accession: S78476; S31211
R/Trueb, B.
submitted to the EMBL Data Library, January 1993
A/Reference number: S78476
A/Accession: S78476
A/Molecule type: mRNA
A/Residues: 1-1888 <TRU>
A/Cross-references: UNIPROT:P32018; EMBL:X70793; NID:G288872; PIDN:CA50064.1; PID:G288875
R/Maechli, C.; Trueb, U.; Kessler, B.; Winterhalter, K.H.; Trueb, B.
Eur. J. Biochem. 212, 483-490, 1993
A/Title: Complete primary structure of chicken collagen XIV.
A/Reference number: S31211; MUID:93185668; PMID:8444186
A/Accession: S31211
A/Structure: preliminary
A/Molecule type: mRNA

A/Residues: 1-416;1460-1811,1843-1888 <WAE>
A/Cross-references: EMBL:X70793
C/Genetics:
A/Gene: Coll14A1
C/Keywords: alternative splicing; coiled coil; extracellular matrix; glycoprotein; trimer
F/1-28/Domain: signal sequence #status predicted <SIG>
F/29-1888/Product: collagen alpha 1(XIV) chain, long form #status predicted <MAT>
F/29-110/Domain: fibronectin type III repeat homology <FN3A>
F/156-330/Domain: von Willebrand factor type A repeat homology <VWA1>
F/352-433/Domain: fibronectin type III repeat homology <FN3B>
F/442-525/Domain: fibronectin type III repeat homology <FN3C>
F/534-614/Domain: fibronectin type III repeat homology <FN3D>
F/623-707/Domain: fibronectin type III repeat homology <FN3E>
F/741-823/Domain: fibronectin type III repeat homology <FN3F>
F/832-914/Domain: fibronectin type III repeat homology <FN3G>
F/922-1009/Domain: fibronectin type III repeat homology <FN3H>
F/1040-1205/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 8.5%; Score 130; DB 2; Length 1888;
Best Local Similarity 25.0%; Pred. No. 0.057; Indels 26; Gaps 12;
Matches 62; Conservative 44; Mismatches 116;
Qy 17 DLVFLDKSGSV-LHHNNEIYFVEQL-AHKFISPO-LRMSFIVSTRGTLMKLT--E 70
Db 1042 DLVFLVDSGWSIGDDNFNKIISFLYSTVGALDKIGDGTQVAILQFSDDPRTFKLNAYK 1101
Qy 71 DREOIRQGLEBLOKVLPGDVTYMEGFERASEQIYYENRQGYRTA--SVIALLDGELEH 128
Db 1102 TKETLLBAIQOI--AVKGGNTKTGKAIKHARE-VLFTGEAGMKRGIPKVLVITDGRSQD 1158
Qy 129 DLFFYSERANRNDL-GAIVYCVGVKDPNFTQALRIAD--SKDHYFPVNDGQALQGI 185
Db 1159 DV-----NKVSRMQLDGSPFAIGVADADYSELVNISSKPSRRHVPFVD--FDAFTKIE 1212
Qy 186 HSILKSCIEILAEPTICAGESFQVNVVRNGFRHARNDVRLCSFKINDSVTLNEKPF 245
Db 1213 DELITFVETASATCPVLVFKDGDKLA-----GFMGMMPGLVKEKFSALDGVSMERGTF 1266
Qy 246 SVEDTYLL 253
Db 1267 NVYPCYRL 1274

RESULT 11
S42373
hypothetical protein T20G5.3 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 07-Oct-1994 #sequence_revision 10-Nov-1995 #text_change 09-May-2004
C/Accession: S42373
R/Smith, A.
submitted to the EMBL Data Library, March 1994
A/Reference number: S42373
A/Accession: S42373
A/Molecule type: DNA
A/Residues: 1-3051 <SMI>
A/Cross-references: EMBL:Z30423; NID:G458479; PID:G458485
C/Genetics:
A/Insertions: 96/1; 166/3; 201/2; 355/1; 404/1; 427/2; 794/1; 892/2; 1041/1; 1146/1;
F/512-679/Domain: von Willebrand factor type A repeat homology <VWA1>
F/754-793/Domain: fibronectin type II repeat homology <EGR>
F/1201-1244/Domain: EGF homology <EGF>

Query Match 8.4%; Score 128; DB 2; Length 3051;
Best Local Similarity 27.1%; Pred. No. 0.15; Indels 66; Gaps 13;
Matches 64; Conservative 31; Mismatches 75;
Qy 12 CYGGF-----DLVFLDKSGSVLHH--NNEIYFVEQLA 43
Db 483 CYGGFVDVSSNANLPPGRVCTVQTCRKQKTDLVFLDGSIGSYVFKNEVLRFAREFV 542
Qy 44 HKFI--SPQLRMSFIVSTRGTLMKLT--DREOIRQGLEBLOKVLPG-----GDTYMH- 94
Db 543 ELFEIRGSKTRVGLIYSDQIRHFPDLDYGDGDSLLKGISFTQ-YLTGLTRGAAIQHM 601

Qy 95 --EGFRASRQIYENRQYR-----TASVITALTDELHEDLPFYSEREANSRDLGAI 147
 Db 602 VQEGF-----SERRGARPOQSDIARVAIILTDGRSQDNV-----TGPADSAKRLSIN 648
 Qy 148 VYCVGVKD-FNETOLARIADSKDHPVNVDSGFQALQGIHSITLKSCTEIIIAEFS 202
 Db 649 TPALGVTHVLASLESLISLGSNNRWFYV-DKFKDLDTLRSMIOK-----AACPS 697

RESULT 12

type XII collagen alpha-1 chain - eastern newt (fragment)
 C/Species: Notophthalmus viridescens, Triturus viridescens (eastern newt)
 C/Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
 C/Accession: 151027
 R/Met, Y./Yang, E.V.; Klatte, K.P.; Tassava, R.A.
 Dev. Biol. 168, 503-513, 1995
 A/Title: Monoclonal antibody MT2 identifies the urodele alpha 1 chain of type XII collagen
 A/Reference number: 151027; MUID:95246925; PMID:7729585
 A/Accession: 151027
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-929 <MBI>
 A/Cross-references: UNIPROT:091145; EMBL:U9494; NID:g632647; PIDN:AAA80217.1; PID:g6326
 F155-236/Domain: Fibronectin type III repeat homology <3FR>
 F1631-795/Domain: von Willebrand factor type A repeat homology <WMA3>

Query Match 7.7%; Score 118; DB 2; Length 929;
 Best Local Similarity 21.2%; Pred. No. 0.21; Mismatches 135; Indels 26; Gaps 9;
 Matches 57; Conservative 51;

Qy 17 DLVYILDKSGV-LHMNEIYFVEQLAHKFIISPOLRMSFTVSTRTGLTKLTEDRQI 75
 Db 633 DIVLLVDSWMSIGRNFKIVRFRVVEVPDGSDRQIVASQVSGPRTMQLNTHKT 692
 Qy 76 RQGLEBLQKLP--GGDTYMEGFERASEQIYENRQY---RTASVITALTDELHED 129
 Db 693 KKSIMDAVANLPYKGGNTNTSGALFIIENNF---RPVGMREKARKIAIILTDGKSODD 749
 Qy 130 LFPYSERANSRDLGAIYCVGVDPFNETOLARIADSKDHPVNV--DGFQALQGIHSI 188
 Db 750 IVASRSKRIA---DEGILYAVGINADENELKEIASPPDELYMTNVADFSULTNIVDL 805
 Qy 189 LKKSCTEIIIAA---EPSTICAGESFQVNVGNGFRHARNYDRLCSFKINDSVTLNEKP 244
 Db 806 TENVCNSVKGPGGLNPENLVTSEPTPRSPFTVTPPSQSYER----FKVEYYPVAGSRP 861
 Qy 245 FSVEDTYLCPAPILKEVGMKALQVSN 273
 Db 862 --QEVYVRGTQTTVLVLGLKRETEYYN 887

RESULT 13

cartilage matrix protein precursor - human
 C/Species: Homo sapiens (man)
 C/Date: 12-Jul-1991 #sequence_revision 12-Jul-1991 #text_change 09-Jul-2004
 C/Accession: A37979; B37979
 R/Jenkins, R.N.; Osborne-Lawrence, S.L.; Sinclair, A.K.; Eddy Jr., R.L.; Byers, M.G.; Sh
 J. Biol. Chem. 265, 19624-19631, 1990
 A/Title: Structure and chromosomal location of the human gene encoding cartilage matrix
 A/Reference number: A37979; MUID:91060568; PMID:2246248
 A/Accession: A37979
 A/Molecule type: DNA
 A/Residues: 1-496 <JEN>
 A/Cross-references: UNIPROT:P21941; GB:J05667
 A/Accession: B37979
 A/Molecule type: mRNA
 A/Residues: 157-290, 'L', 292-496 <JEN>
 A/Cross-references: GB:M55683; GB:J05666; GB:J05667; NID:g180651; PIDN:AAA63904.1; PID:g
 C/Accession: A37979
 A/Title: Amino acid sequence of the Bb fragment from complement factor B. Sequence of the

A/Cross-references: GDB:127280; OMIM:115437
 A/Map position: 1p35-1p35
 A/Introns: 32/1; 147/3; 222/1; 264/1; 403/1; 454/1; 481/1
 C/Complex: homotrimer
 C/Superfamily: cartilage matrix protein; EGF homology; von Willebrand factor type A repe
 C/Keywords: glycoprotein, homotrimer
 F11-42/Domain: signal sequence #status predicted <SIG>
 F13-496/Product: cartilage matrix protein #status predicted <MAT>
 F139-206/Domain: von Willebrand factor type A repeat homology <WMA1>
 F127-262/Domain: EGF homology <EGF>
 F127-437/Domain: von Willebrand factor type A repeat homology <WMA2>
 F176-344/Binding site: carbohydrate (asn) (covalent) #status predicted
 F1221-228, 234-247, 249-262/Dissulfide bonds: #status predicted

Query Match 7.4%; Score 113.5; DB 2; Length 496;
 Best Local Similarity 26.0%; Pred. No. 0.21;
 Matches 59; Conservative 35; Mismatches 92; Indels 41; Gaps 13;

Qy 12 CYGG----FDLYILDKSGV-L-HMNEIYFVEQLAHKFIISPOLRMSFTV--FSTR 61
 Db 265 CSGCGSSATDVLVFLIDGSKSVDPENFELVKFISQIVDTLVDKLAQVGLVQSSSVR 324
 Qy 62 GTTLMKLTEDREQRQGLEBLQKLP--GGDTYMEGFERASEQIY---YENRQYRTAS 116
 Db 325 QEPFLGRFHTKDKIAVNRN-----SYMEKGTWTGALKYLIINSFTVSSGARRPA 376
 Qy 117 --VITALTDELHEDLPFYSEREANSRDLGAIYCVGVDPFNETOLARIADS--KDHVF 172
 Db 377 QKGVIVTDRSQD---YINDAKKAKDGFKNFANGVNAVDRLREINSEVAHYF 432
 Qy 173 PVNDGFQALQGIHSITLKSCTEIIIAEPSTICAGES--FQVNVG 216
 Db 433 VTAD-FRTINQIGKKLQKICV---BDP---CACESLVFQAKVEG 472

RESULT 14

BBHU
 Complement factor B precursor [validated] - human
 N/Alternate names: C3 convertase; C3 proactivator; glycine-rich beta-glycoprotein; heat-j
 N/Contents: alternative-complement-pathway C3/C5 convertase (EC 3.4.21.47) Bb fragment
 C/Species: Homo sapiens (man)
 C/Date: 19-Feb-1984 #sequence_revision 05-Aug-1994 #text_change 09-Jul-2004
 C/Accession: S34075; A44622; A00934; A19188; A19947; B19947; B25971; S14339; A44628; I544
 R/Mejia, J.E.; Jahn, I.; de la Saille, H.; Hauptmann, G.
 submitted to the EMBL Data Library, March 1993
 A/Reference number: S34075
 A/Accession: S34075
 A/Molecule type: mRNA
 A/Residues: 1-764 <MEJ>
 A/Cross-references: UNIPROT:P00751; EMBL:X72875; NID:g297568; PIDN:CAA51389.1; PID:g29756
 R/Woods, D.E.; Markham, A.F.; Ricker, A.T.; Goldberger, G.; Colten, H.R.
 Proc. Natl. Acad. Sci. U.S.A. 79, 5661-5665, 1982
 A/Title: Isolation of cDNA clones for the human complement protein factor B, a class III
 A/Reference number: A44622; MUID:83039428; PMID:6957884
 A/Accession: A44622
 A/Molecule type: DNA
 A/Residues: 467-546; 550-595; 752-764 <WOO>
 A/Cross-references: GB:J00185; GB:J00186
 A/Note: the authors translated the codon TAC at 519 as Thr; the nucleic acid translation
 R/Moile, J.B.; Anderson, J.K.; Davison, E.A.; Woods, D.E.
 J. Biol. Chem. 259, 3407-3412, 1984
 A/Title: Complete primary structure for the zymogen of human complement factor B.
 A/Reference number: A20751; MUID:84161997; PMID:6546754
 A/Accession: A00934
 A/Molecule type: protein; mRNA
 A/Residues: 26-764 <MOI>
 A/Cross-references: GB:K01566
 A/Note: nucleic acid translation differs from the sequence shown in having 300-Leu, 328-V
 A/Note: glycosylation sites were determined
 R/Christie, D.L.; Gagnon, J.
 Biochem. J. 209, 61-70, 1983
 A/Title: Amino acid sequence of the Bb fragment from complement factor B. Sequence of the

A:Reference number: A19188; MUID:83204002; PMID:6342610
A:Contents: the final paper in a series documenting the sequence, glycosylation site, and
A:Accession: A19188
A:Molecule type: protein
A:Residues: 260-296, 'T', 298-764 <CHR>
R:Campbell, R.D.; Porter, R.R.
Proc. Natl. Acad. Sci. U.S.A. 80, 4464-4468, 1983
A:Title: Molecular cloning and characterization of the gene coding for human complement
A:Reference number: A19947; MUID:83273641; PMID:6308626
A:Accession: A19947
A:Molecule type: DNA
A:Residues: 346-764 <CAM>
A:Cross-references: GB:J00125
A:Accession: B19947
A:Molecule type: mRNA
A:Residues: 339-509 <CA1>
A:Cross-references: GB:J00126; NID:G187723; PIDN:AAA6226.1; PID:G553536
R:Mu, L.; Morley, B.J.; Campbell, R.D.
Cell 48, 331-342, 1987
A:Title: Cell-specific expression of the human complement protein factor B gene: evidence
A:Reference number: A25971; MUID:87102880; PMID:3643061
A:Accession: B25971
A:Molecule type: DNA
A:Residues: 1-99 <WLU>
A:Cross-references: GB:M5082; NID:G187693; PIDN:AAA59625.1; PID:G553534
R:Niemann, M.A.; Brown, A.S.; Miller, E.J.
Biochem. J. 274, 473-480, 1991
A:Title: The principal site of glycation of human complement factor B.
A:Reference number: S14339; MUID:91174758; PMID:2006911
A:Accession: S14339
A:Molecule type: protein
A:Residues: 270-329 <IE>
A:Note: binding site for carbohydrate to lysine under artificial conditions
R:Morley, B.J.; Campbell, R.D.
EMBO J. 3, 153-157, 1984
A:Title: Internal homologues of the Ba fragment from human complement component factor B
A:Reference number: A44620; MUID:84158524; PMID:6323161
A:Accession: A44628
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 16-225, 'P', 227-259 <MOR>
R:Schwaeble, W.; Luttig, B.; Sokolowski, T.; Esteller, C.; Weiss, E.H.; Meyer zum Busche
Immunobiology 188, 221-232, 1993
A:Title: Human complement factor B: functional properties of a recombinant zymogen of the
A:Reference number: I54409; MUID:94041399; PMID:8225386
A:Accession: I54409
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-764 <RES>
A:Cross-references: GB:S67310; NID:G452937; PIDN:ADJ1389.1; PID:G4261689
R:Horvath, T.; Klm, S.; Matsumoto, M.; Watanabe, I.; Fujita, S.; Volanakis, J.E.
Mol. Immunol. 30, 1587-1592, 1993
A:Title: Human complement factor B: cDNA cloning, nucleotide sequencing, phenotypic conve
A:Reference number: I57824; MUID:94067177; PMID:8247029
A:Accession: I57824
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-31, 'O', 33-764 <RE2>
A:Cross-references: GB:L15702; NID:G291921; PIDN:AAA16620.1; PID:G251922
C:Comment: 292-Cys has a free sulfhydryl.
C:Genetics:
A:Gene: GDB:BF
A:Cross-references: GDB:119726; OMIM:138470
A:Map position: 6p21.3-6p21.3
A:Introns: 21/3; 99/3; 346/1; 390/1; 424/1; 470/1; 502/3; 542/1; 593/2; 619/1; 652/3; 65
A:Note: the list of introns may be incomplete
C:Complex: complement factor B initially forms an inactive complex with complement fact
ment factor C3b forming active C3/C5 convertase; Ba is released
C:Function:
A:Description: Bb is a serine protease; C3/C5 convertase cleaves complement C3 alpha c
at

C:Superfamily: complement B/2; complement factor H repeat homology; trypsin homology; VC
 C:Keywords: acute phase; complement alternate pathway; duplication; glycoprotein; hydrolytic
 F.1-25/Domain: signal sequence #status predicted <Sig>
 F.26-764/Product: complement factor B #status experimental <MAT>
 F.26-764/Product: complement factor B #status experimental <BAF>
 F.26-764/Product: complement factor B #status experimental <FHI>
 F.37-98/Domain: complement factor H repeat homology <FHI>
 F.103-158/Domain: complement factor H repeat homology <FHI>
 F.165-218/Domain: complement factor H repeat homology <FHI>
 F.260-764/Product: C3/C5 convertase Bb fragment #status experimental <BBF>
 F.268-458/Domain: von Willebrand factor type A repeat homology <VFA>
 F.482-752/Domain: trypsin homology #status atypical <TRY>
 F.37-76, 62-98, 103-145, 131-158, 165-205, 191-218, 478-596, 511-527, 599-615, 656-682, 695-725/Di
 F.122, 142, 285, 378/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F.250-260/Cleavage site: Arg-Lys (covalent) factor D) #status experimental
 F.526, 576, 659/Active site: His, Asp, Ser #status experimental

Query Match 7.4%; Score 113.5; DB 1; Length 764;
 Best Local Similarity 19.6%; Pred. No. 0.37;
 Matches 66; Conservative 63; Mismatches 109; Indels 99; Gaps 17;

QY 1 QCGRRDGG-GRP-----CYGSEFDLYFLDKSGSV-----LHMMETIYFVEQLAH 44
 DB 243 EGVNDEDDGGHPGPOOKRKIVLDPGSGMNIYLVLDGSDISGASNFTGAKKCLVNIIEKVAS 302
 QY 45 KFISQLMNSFIVFSTRTGTLTKLTPE---DPEQIQSGEEL---QKVLPGDGYWHEG 96
 DB 303 YGVKPK--KGLVTVATYPIKIMWKVSEADSSNDMWTKOLNEIYEDHKLSGTVT----- 355
 QY 97 FERASEQIYENR-----QGV-RTASVITIALTDS-----ELHEDLPFVSE 135
 DB 356 --KKAQAVYSMWSWPDVPEEGMNRTRHYIILMTDGLHNMGGPITVYIDIRDLIYIGKD 414
 QY 136 REANFSRDLGAIVCYG--VKDPNETQLARLADSKDHPVPVNDGFOALQGIHSHILKSC 193
 DB 415 RKNPREDYLDVYVFGGPIVNOVNIALASKKONEQHFVKYD--MENTLBOVFYQIMDES- 472
 QY 194 IEILAEPTICAGESFOVYVANGNGFRHARNDVRLCSFKINSVTLNEKPSVEDRYLL 253
 DB 473 -----QSLSLC-----GMVHRKRGTD-----YHKQPMQAKISV-- 501
 QY 254 CPAPIKEVGMKRALQVSMNDGLSFSSVVIITTC 290
 DB 502 ----IRPSKGHESCMG-----AVVSEYFVLTAAHC 527

RESULT 15
 A54849
 collagen alpha 1(VII) chain precursor - human
 M:Alternate names: procollagen alpha 1(VII) chain
 C:Species: Homo sapiens (man)
 C:Date: 04-Nov-1994 #sequence revision 04-Nov-1994 #text change 09-Jul-2004
 C:Accession: A54849; GI61316; I56328; J30296; I84686
 R:Christiano, A.M.; Greenspan, D.S.; Lee, S.; J.
 J. Biol. Chem. 269, 20256-20262, 1994
 A:Title: Cloning of human type VII collagen. Complete primary sequence of the alpha1(VII)
 A:Reference number: A54849; MUID:94327588; PMID:8051117
 A:Accession: A54849
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-2944 <CHR>
 A:Cross-references: UNIPROT:Q02388; GB:L02870; NID:g987124; PIDD:AAA75438.1; PID:g987125
 R:Tanaka, T.; Takahashi, K.; Furukawa, F.; Imanura, S.
 Biochem. Biophys. Res. Commun. 183, 958-963, 1992
 A:Title: Molecular cloning and characterization of type VII collagen cDNA.
 A:Reference number: PH0844; MUID:92231902; PMID:1567409
 A:Accession: PH0844
 A:Molecule type: mRNA
 A:Residues: 'EPR', 340-475, 'RALSTASHSTLCRARTRWECNKGSHWTRACAEFCNRPASHRAARAG', 524-528, 'C', 5, 'S', 15
 A:Cross-references: DDBJ:D11152; DDBJ:D13694; NID:g453698; PIDD:BA02853.1; PID:g453699
 A:Experimental source: Keratinocyte
 A:Note: The authors translated the codon AOC for residues 394 and 397 as Tyr
 R:Patente, M.G.; Chung, L.C.; Ryyanen, J.; Woodley, D.T.; Wynn, K.C.; Bauer, E.A.; Matris
 Proc. Natl. Acad. Sci. U.S.A. 88, 6931-6935, 1991

A>Title: Human type VII collagen: cDNA cloning and chromosomal mapping of the gene.
 A/Reference number: S6316; MUID:91334380; PMID:1871109
 A/Accession: S16316
 A/Molecule type: mRNA
 A/Residues: 815-892, 'E', 894-1439 <PAR>
 A/Cross-references: GB:M65158; GB:649017; NID:G180914; PIDN:AAA96439.1; PID:G180915
 A/Experimental source: keratinocyte
 R/Gammon, W.R.; Abernethy, M.L.; Pedilla, K.M.; Pilsayanh, P.S.; Cook, M.E.; Wright, J.J.
 J. Invest. Dermatol. 99, 691-696, 1992
 A>Title: Noncollagenous (NC1) domain of collagen VII resembles multidomain adhesion prot
 A/Reference number: 156328; MUID:93107742; PMID:1469284
 A/Accession: 156328
 A/Status: translated from GB/EMBL/DBU
 A/Molecule type: mRNA
 A/Residues: 'ERR', 372-517, 'DV', 520-540, 'W', 542-1255 <RES>
 A/Cross-references: GB:S51236; NID:G262308; PIDN:AB24637.1; PID:G262309
 R/Selbzer, J.L.; Eisen, A.Z.; Bauer, E.A.; Morris, N.P.; Gnanville, R.W.; Burgeson, R.E.
 J. Biol. Chem. 264, 3822-3826, 1989
 A>Title: Cleavage of type VII collagen by interstitial collagenase and type IV collagenase
 A/Reference number: A30296; MUID:89139437; PMID:2537292
 A/Accession: A30296
 A/Molecule type: Protein
 A/Residues: 'A', 1240-1246, 'G', 1248-1250, 'XE', 1253-1255, 'Q', 1257, 'B', 2032, 'C', 2034-2041, '
 A/Note: two reported peptide sequences cannot be reliably located
 R/Greenspan, D.S.
 Hum. Mol. Genet. 2, 273-278, 1993
 A>Title: The carboxyl-terminal half of type VII collagen, including the non-collagenous
 A/Reference number: 148103; MUID:93271985; PMID:8499916
 A/Accession: 148103
 A/Status: preliminary; translated from GB/EMBL/DBU
 A/Molecule type: mRNA
 A/Residues: 2395-2871, 'S', 2873-2944 <RE2>
 A/Cross-references: GB:L06862; NID:G388713; PIDN:AAA89196.1; PID:G388714
 R/Christiano, A.M.; Rymaszewski, M.; Uitto, J.
 Proc. Natl. Acad. Sci. U.S.A. 91, 3549-3553, 1994
 A>Title: Dominant dystrophic epidermolysis bullosa: identification of a Gly --- Ser substitution
 A/Reference number: A55255; MUID:94224777; PMID:8170945
 A/Accession: A55255
 A/Comment: Prolines and lysines at the third position of the tripeptide repeating unit (C
 ed and subsequently O-glycosylated.
 C/Genetics:
 A/Gene: GDB:COL7A1; EBR1; EBD1; EB
 A/Cross-references: GDB:128750; OMIM:120120
 A/Map position: 3p21.3-3p21.3
 A/Note: defects in this gene can result in dominant and recessive dystrophic epidermolysis
 A/Note: there are 118 introns
 C/Complex: type VII collagen is probably a homotrimer
 C/Function:
 A/Description: structural component of extracellular polymer associated with anchoring f
 C/Keywords: coll1; extracellular matrix; glycoprotein; hydroxylysine; hydroxyprolin
 F/1-16/Domain: signal sequence #status predicted <SIG>
 F/17-2944/Product: collagen alpha 1 (VII) chain #status predicted <MAT>
 F/17-1253/Domain: amino-terminal nonhelical #status predicted <NC1>
 F/36-201/Domain: von Willebrand factor type A repeat homology <VWA1>
 F/231-318/Domain: fibronectin type III repeat homology <FN1>
 F/327-413/Domain: fibronectin type III repeat homology <FN2>
 F/414-502/Domain: fibronectin type III repeat homology <FN3>
 F/508-593/Domain: fibronectin type III repeat homology <FN4>
 F/598-683/Domain: fibronectin type III repeat homology <FN5>
 F/686-771/Domain: fibronectin type III repeat homology <FN6>
 F/776-862/Domain: fibronectin type III repeat homology <FN7>
 F/864-952/Domain: fibronectin type III repeat homology <FN8>
 F/954-1045/Domain: fibronectin type III repeat homology <FN9>
 F/1052-1219/Domain: von Willebrand factor type A repeat homology <VWA2>
 F/1170-1172/Region: cell attachment (R-G-D) motif
 F/1189-1253/Region: cysteine/proline-rich
 F/1254-2783/Region: interrupted helical
 F/1334-1336/Region: cell attachment (R-G-D) motif
 F/2008-2010/Region: cell attachment (R-G-D) motif
 F/2553-2555/Region: cell attachment (R-G-D) motif
 F/2784-2944/Domain: carboxyl-terminal nonhelical #status predicted <NC2>
 F/2876-2929/Domain: animal Kunitz-type proteinase inhibitor homology <SPI>
 F/337,786,1109/Binding site: carbohydrate (Aen) (covalent) #status predicted

F/2167,2176,2185,2188,2664,2667,2673/Modified site: 4-hydroxyproline (Pro) #status exper:
 F/2625,2631/Modified site: 5-hydroxylysine (Lys) #status experimental
 F/2625,2631/Binding site: carbohydrate (Lys) (covalent) #status experimental
 F/2634,2802,2804/Disulfide bonds: interchain #status predicted

Query Match 7.38; Score 111.5; DB 2; Length 2944;
 Best Local Similarity 24.18; Pred. No. 3.2;
 Matches 52; Conservative 42; Mismatches 85; Indels 37; Gaps 11;
 17 DLYEILDKSGSV-LHNMNEIYFVEOLAHKFE---ISPO-LRMSPIVSTRGTT---LMKL 68
 38 DIVFLLGSSSIGRSNREVRSLFEGVLVLPFGAASQGVRAFTVQYSDPPRIFEGDAL 97
 69 TEDREIQIGLIELQKVLPGSDTYMEGFERRASBOIYENROGYRTASVITALTDEGLIE 128
 98 GSGGDVIR-AIRBELS--YKGNRTGAALIHVADHVLPLQALPAGVKKVCIILTDGR-SQ 153
 129 DLPFYSREANRSHDLCATYICGVKDFNETQLARLD--SKDHVPVNDGFOALQGIH 186
 154 DLV---DTAAQRUKGQGVKLFVAVGINKADPEELKRVASQPTSDFFFEVND-FSILRTLLP 209
 187 SILKSGCI-----EILARSPST 203
 210 LVSRVCTTAGVVPVTRPPDDSTAPRDLVLSPPS 245

Search completed: June 13, 2005, 20:06:35
 Job time : 21.0213 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using BW model

Run on: June 13, 2005, 19:37:47 ; Search time 98.3587 Seconds

(without alignments)
1525.429 Million cell updates/sec

Title: US-09-970-076-2_COPY_28_320

Perfect score: 1526

Sequence: 1 QGGRREDGACGACGPDLYF.....DGLSFSSSVITTTTCSGDG 293

Scoring table: BLOSUM62

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1526	100.0	564	1 ATR1_HUMAN	Q9H6X2 homo sapien
2	1509	98.9	562	1 ATR1_MOUSE	Q9CZ52 mus musculu
3	786	51.5	489	1 ATR2_HUMAN	P58J35 homo sapien
4	770	50.5	487	2 Q6DFX2	Q6DFX2 mus musculu
5	606	39.7	641	2 Q8BVM2	Q8BVM2 mus musculu
6	388	25.4	97	2 Q9BEC6	Q9BEC6 homo sapien
7	159.5	10.5	1161	1 ITAD_RAT	Q9GYE7 rattus norv
8	156	10.2	1162	1 ITAD_HUMAN	Q13349 homo sapien
9	153.5	10.1	1332	2 Q9BPQ8	Q9BPQ8 halocynthia
10	148.5	9.7	1626	2 Q8NFW1	Q8NFW1 homo sapien
11	148	9.7	441	2 Q8T6U5	Q8T6U5 mytilus edu
12	148	9.7	444	2 Q8T5C3	Q8T5C3 mytilus edu
13	145.5	9.5	1152	1 ITAM_HUMAN	P11215 homo sapien
14	145	9.5	453	2 Q8T5C2	Q8T5C2 mytilus gal
15	143.5	9.4	1163	1 ITAX_HUMAN	P20702 homo sapien
16	143	9.4	3124	1 CA1C_CHICK	P13944 gallus galli
17	141	9.2	1153	1 ITAM_MOUSE	P05555 mus musculu
18	139	9.1	724	2 Q04588	Q04588 elmeria max
19	139	9.1	3119	1 CA1C_MOUSE	Q04588 elmeria ten
20	134.5	8.8	712	2 Q03981	Q03981 elmeria ten
21	134	8.8	760	2 Q070350	Q070350 mus musculu
22	132.5	8.7	765	2 Q9U8U5	Q9U8U5 neospora ca
23	131.5	8.6	517	2 Q03853	Q03853 homo sapien
24	131	8.6	760	1 CO2_MOUSE	P21180 mus musculu
25	131	8.6	3063	1 CA1C_HUMAN	Q99715 homo sapien
26	130.5	8.6	920	2 Q28984	Q28984 sus scrofa
27	130	8.5	1888	1 CA1E_CHICK	P32018 gallus galli
28	129.5	8.5	1182	2 Q8C6F9	Q8C6F9 mus musculu
29	128	8.4	221	2 Q99L17	Q99L17 mus musculu
30	128	8.4	3183	2 Q65ZC2	Q65ZC2 caenorhabdi
31	128	8.4	3767	1 MOA3_CAEBL	P34576 caenorhabdi

32	123	8.1	1151	2 Q9J130	Q9J130 rattus norv
33	121	7.9	758	2 Q6MG73	Q6MG73 rattus norv
34	121	7.9	758	2 Q8CIP8	Q8CIP8 rattus norv
35	121	7.9	2401	2 Q7RF52	Q7RF52 plasmodium
36	120.5	7.9	790	2 Q6DC06	Q6DC06 xenopus lae
37	118.5	7.8	440	2 Q8CBT2	Q8CBT2 mus musculu
38	118.5	7.8	848	2 Q8C720	Q8C720 mus musculu
39	118.5	7.8	919	2 Q7SR52	Q7SR52 lymanaea sta
40	118.5	7.8	3567	2 Q9ES77	Q9ES77 mus musculu
41	118	7.7	652	2 Q9SL12	Q9SL12 bos taurus
42	118	7.7	929	1 CA1C_NOTVI	Q9145 notophthalm
43	117.5	7.7	637	2 Q8IVX1	Q8IVX1 homo sapien
44	117.5	7.7	1284	2 Q6P159	Q6P159 homo sapien
45	117.5	7.7	1329	1 KF10_HUMAN	Q9P218 homo sapien

ALIGNMENTS

RESULT 1
ID ATR1_HUMAN STANDARD: PRT; 564 AA.
AC Q9H6X2: Q96P02; Q9NVP3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Anthrax toxin receptor 1 precursor (Tumor endothelial marker 8).
GN Name=ANTXR1; Synonyms=ATR, TEM8;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=20407466; PubMed=10947988; DOI=10.1126/science.289.5482.1197;
RA St Croix B., Rago C., Velculescu V., Traverso G., Romans K.B.,
RA Montgomery E., Lai A., Riggs G.J., Lengauer C., Vogelstein B.,
RA Kinzler K.W.;
RT "Genes expressed in human tumor endothelium";
RL Science 289:1197-1202(2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2), AND INTERACTION WITH ANTHRAX TOXIN.
RX MEDLINE=21557240; PubMed=11700562; DOI=10.1038/n35101998;
RA Bradley K.A., Mogridge J., Mourer M., Collier R.J., Young J.A.T.;
RT "Identification of the cellular receptor for anthrax toxin.";
RL Nature 414:225-229(2001).
RN [3]
RP SEQUENCE OF 184-564 FROM N.A. (ISOFORM 1), AND SEQUENCE FROM N.A.
RP (ISOFORM 3).
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Ohtsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Nagahori K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuna M.,
RA Shiratori A., Sudo H., Hosioki T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikawa E.,
RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RA Yamazaki M., Nishimura K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoaka S., Chiba Y.,
RA Iehida S., Ono Y., Takiguchi S., Watanabe S., Yoshida M., Hotsu T.,
RA Kuwano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Togiyasu S., Komai F., Hara R., Takeuchi K., Arita M.,
RA Imose N., Masehino K., Yuki H., Oshima A., Saeki N., Aosekura S.,
RA Yoshihawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishiyaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki Y., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujikawa T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senda T.,

RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Maeno Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isegai T., Sugano S.,
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RT Nat. Genet. 36:40-45 (2004).
RN [4]
RN SEQUENCE FROM N.A. (ISOFORM 4).
RC TISSUE=Kidney; PubMed=12477932; DOI=10.1073/pnas.242603899;
RX MEDLINE=22386257; Feingold E.A., Grouse L.H., Derge J.G.,
RX Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Schaefer C.F., Bhat N.K.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Ditchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stopleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carrincci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield V.S.N., Krzywinski M.I., Skalska J., Smalins D.E.,
RA Scherch A., Schin J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [5]
RN INTERACTION WITH ANTHRAX TOXIN.
RP TISSUE=Placenta; PubMed=12700348; DOI=10.1073/pnas.0431098100;
RX MEDLINE=22608610; Bradley K.A., Young J.A.T.,
RX Scobie H.M., Ratney G.J.A., Bradley K.A., Young J.A.T.,
RT "Human capillary morphogenesis protein 2 functions as an anthrax toxin
RT receptor.";
RT Proc. Natl. Acad. Sci. U.S.A. 100:5170-5174 (2003).
RN [6]
RN SPLICE ISOFORM(S) THAT ARE POTENTIAL NMD TARGET(S).
RP PubMed=14759258; DOI=10.1186/gb-2004-5-2-r8;
RX Hillman R.T., Green R.E., Brenner S.B.,
RA "An unappreciated role for RNA surveillance.";
RT Genome Biol. 5:RESEARCH008.1-RESEARCH008.16(2004).
CC -1- FUNCTION: Cellular role is not yet known.
CC -1- SUBUNIT: Binds to the protective antigen (PA) of *Bacillus anthracis*. Binding does not occur in the presence of calcium.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Probable).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Comment=Experimental confirmation may be lacking for some
CC isoforms;
CC Name=1;
CC IsoId=Q9H6X2-1; Sequence=Displayed;
CC Note=May be produced at very low levels due to a premature stop
CC codon in the mRNA, leading to nonsense-mediated mRNA decay;
CC Name=2;
CC IsoId=Q9H6X2-2; Sequence=VSP_000444, VSP_000445;
CC Name=3;
CC IsoId=Q9H6X2-3; Sequence=VSP_000446, VSP_000447;
CC Name=4;
CC IsoId=Q9H6X2-4; Sequence=VSP_000448, VSP_000449;
CC -1- TISSUE SPECIFICITY: Highly expressed in tumor endothelial cells
CC but not in normal endothelial cells.
CC -1- DOMAIN: Binding to PA seems to be effected through the VWA domain.
CC -1- SIMILARITY: Belongs to the ATR family.
CC -1- SIMILARITY: Contains 1 VWA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL: AF279145; AKS2094.1; -;
CC DR EMBL: AF421380; AAL26496.1; -;
CC DR EMBL: AK025429; BAB15128.1; ALT_INIT.
CC DR EMBL: AK001463; BAA91707.1; ALT_FRAME.
CC DR EMBL: BC012074; AAI12074.1; -;
CC DR Gene: HGNC:21014; ANTXR1.
CC H-InvDB: HIX0002125; -;
CC MIM: 606410; -;
CC Interpro: IPR008400; Anth_Ig.
CC Interpro: IPR008399; Ant_C.
CC Interpro: IPR002035; VWF_A.
CC Pfam: PF05587; Anth_Ig; 1.
CC Pfam: PF05586; Ant_C; 1.
CC Pfam: PF00092; VWA; 1.
CC SMART: SM00327; VWA; 1.
CC POSITIVE: PS50234; VWA; 1.
CC KW Alternative splicing; Glycoprotein; Receptor; Signal; Transmembrane.
CC STGNL 1
CC FT CHAIN 33
CC FT DOMAIN 33
CC FT TRANSSEM 322
CC FT DOMAIN 343
CC FT DOMAIN 44
CC FT DOMAIN 360
CC FT DOMAIN 506
CC FT CARBOHYD 166
CC FT CARBOHYD 184
CC FT CARBOHYD 262
CC FT VARSPPLIC 365
CC FT VARSPPLIC 369
CC FT VARSPPLIC 268
CC FT VARSPPLIC 298
CC FT VARSPPLIC 319
CC FT VARSPPLIC 334
CC FT SEQUENCE 564 AA; 62789 MW; B118A00D5DF2233 CRC64;
CC Query Match 100.0%; Score 1526; DB 1; Length 564;
CC Best Local Similarity 100.0%; Pred. No. 3,3e-112;
CC Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC Qy 1 QGGRREGGGACGCGPGLYFTLLKSGSVLHHNMEIYFVQVLAKHFTPLRMSTFYEST 60
CC Db 28 QGGRREGGGACGCGPGLYFTLLKSGSVLHHNMEIYFVQVLAKHFTPLRMSTFYEST 87
CC Qy 61 RGTTLMLTLDPRQIRGLELQKVLPGCDTYNHEGGERASEQIYYENRGQYRTASYIA 120
CC Db 88 RGTTLMLTLDPRQIRGLELQKVLPGSDTYNHEGGERASEQIYYENRGQYRTASYIA 147
CC Qy 121 LTQDELHEDLFFYSERERANSRDIGALIVCVKDFENETOLARIADSKDHVPVNDGFOA 180
CC Db 148 LTQDELHEDLFFYSERERANSRDIGALIVCVKDFENETOLARIADSKDHVPVNDGFOA 207
CC Qy 181 LOGTTHILKSKCEIILAAPSTTCAGESQVYVNRGNGPFRHANNRVLCSPKINSVTL 240
CC Db 208 LOGTTHILKSKCEIILAAPSTTCAGESQVYVNRGNGPFRHANNRVLCSPKINSVTL 267
CC Qy 241 NEKPFVEDTYLLCPAPILKEVGKAKALQVSNDSGLFISSVLIITTHCSDG 293
CC Db 268 NEKPFVEDTYLLCPAPILKEVGKAKALQVSNDSGLFISSVLIITTHCSDG 320

```
RESULT 2
A1R1_MOUSE STANDARD; PRT; 562 AA.
ID A1R1_MOUSE
AC Q9C252;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Anthrax toxin receptor 1 precursor (tumor endothelial marker 8).
GN Name=Anthrax1; Synonyms=Atcr, Tem8;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 1;
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=21443268; PubMed=11559528;
RA Carson-Walter E.B., Watkins D.N., Nanda A., Vogelstein B.,
RA Kinzler K.W., St Croix B.;
RT "Cell surface tumor endothelial markers are conserved in mice and
humans."
RL Cancer Res. 61:6649-6655 (2001).
RN 2;
RP SEQUENCE OF 88-562 FROM N.A. (ISOFORM 2).
RX STRAIN=C57BL/6J; TISSUE=Embryo;
MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Boro H., Kondo S.,
RA Nikiel D., Ogaso N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Nagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojohori T.,
RA Baldeirrelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schirni L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chochia C., Corbani L.B., Cousins S.,
RA Dalla E., Drgenti T.A., Fletcher C.F., Forrester A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Giesel C., Godzik A., Gough J.,
RA Grimmond S., Guerinch S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawanishi H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Matsumoto K., Marchionni L., McKenzie L., Miki H.,
RA Nagaishima T., Numa K., Okada T., Pavan W.J., Pereira G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sanderlin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Waples C., Wang Y., Watanabe Y., Wells C.,
RA Wilmberg L.G., Wyshaw-Boris A., Yanagisawa M., Yang L., Yang L.,
RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
CC -1- FUNCTION: Cellular role is not yet known.
CC -1- SUBUNIT: Binds to the protective antigen (PA) of Bacillus
anthracis (by similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Probable).
CC -1- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named Isoforms=2;
Name=1;
CC Name=2;
CC IsoId=Q9C252-1; Sequence=Displayed;
CC Name=3;
CC IsoId=Q9C252-2; Sequence=VSP_000450;
CC Note=No experimental confirmation available;
CC -1- DOMAIN: Binding to PA seems to be effected through the VWA domain
(by similarity).
CC -1- SIMILARITY: Belongs to the ATR family.
CC -1- SIMILARITY: Contains 1 VWA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announce/
CC or send an email to license@ebi.ac.uk).
```

QY	2	GGSRREDGPGACGCGFDLYFLIDKSGSVLHFMNITVYVDELAKFTISPOLRMSPTIVSTR	61
QY <td>27<td>GGRRREDGPGACGCGFDLYFLIDKSGSVLHFMNITVYVDELAKFTISPOLRMSPTIVSTR<td>86</td></td></td>	27 <td>GGRRREDGPGACGCGFDLYFLIDKSGSVLHFMNITVYVDELAKFTISPOLRMSPTIVSTR<td>86</td></td>	GGRRREDGPGACGCGFDLYFLIDKSGSVLHFMNITVYVDELAKFTISPOLRMSPTIVSTR <td>86</td>	86
QY <td>62<td>GTTLMLKTEREQRQGLLEIQLKVLPGCDTYMHEGFEPRASEQIYENRQGYRTASVITIAL<td>121</td></td></td>	62 <td>GTTLMLKTEREQRQGLLEIQLKVLPGCDTYMHEGFEPRASEQIYENRQGYRTASVITIAL<td>121</td></td>	GTTLMLKTEREQRQGLLEIQLKVLPGCDTYMHEGFEPRASEQIYENRQGYRTASVITIAL <td>121</td>	121
QY <td>87<td>GTTLMLKTEREQRQGLLEIQLKVLPGCDTYMHEGFEPRASEQIYENRQGYRTASVITIAL<td>146</td></td></td>	87 <td>GTTLMLKTEREQRQGLLEIQLKVLPGCDTYMHEGFEPRASEQIYENRQGYRTASVITIAL<td>146</td></td>	GTTLMLKTEREQRQGLLEIQLKVLPGCDTYMHEGFEPRASEQIYENRQGYRTASVITIAL <td>146</td>	146
QY <td>122<td>TGDELHEDLFFYSERBRANRSDIGAIYVCVGDPMETQIARIADSKDHYFPVNDGFQAL<td>181</td></td></td>	122 <td>TGDELHEDLFFYSERBRANRSDIGAIYVCVGDPMETQIARIADSKDHYFPVNDGFQAL<td>181</td></td>	TGDELHEDLFFYSERBRANRSDIGAIYVCVGDPMETQIARIADSKDHYFPVNDGFQAL <td>181</td>	181
QY <td>147<td>TGDELHEDLFFYSERBRANRSDIGAIYVCVGDPMETQIARIADSKDHYFPVNDGFQAL<td>206</td></td></td>	147 <td>TGDELHEDLFFYSERBRANRSDIGAIYVCVGDPMETQIARIADSKDHYFPVNDGFQAL<td>206</td></td>	TGDELHEDLFFYSERBRANRSDIGAIYVCVGDPMETQIARIADSKDHYFPVNDGFQAL <td>206</td>	206
QY <td>182<td>QGIHSLTKSCIEIILAAESTICAGSSFOVVRNGGFRARAVDRILGCFKINDSVTLN<td>241</td></td></td>	182 <td>QGIHSLTKSCIEIILAAESTICAGSSFOVVRNGGFRARAVDRILGCFKINDSVTLN<td>241</td></td>	QGIHSLTKSCIEIILAAESTICAGSSFOVVRNGGFRARAVDRILGCFKINDSVTLN <td>241</td>	241
QY <td>207<td>QGIHSLTKSCIEIILAAESTICAGSSFOVVRNGGFRARAVDRILGCFKINDSVTLN<td>266</td></td></td>	207 <td>QGIHSLTKSCIEIILAAESTICAGSSFOVVRNGGFRARAVDRILGCFKINDSVTLN<td>266</td></td>	QGIHSLTKSCIEIILAAESTICAGSSFOVVRNGGFRARAVDRILGCFKINDSVTLN <td>266</td>	266
QY <td>242<td>KEPFSVEDTYLCPAPILKEVGKKAALQVSMNDGLSFISSVITITTHCSDG<td>293</td></td></td>	242 <td>KEPFSVEDTYLCPAPILKEVGKKAALQVSMNDGLSFISSVITITTHCSDG<td>293</td></td>	KEPFSVEDTYLCPAPILKEVGKKAALQVSMNDGLSFISSVITITTHCSDG <td>293</td>	293
QY <td>267<td>KEPFSVEDTYLCPAPILKEVGKKAALQVSMNDGLSFISSVITITTHCSDG<td>318</td></td></td>	267 <td>KEPFSVEDTYLCPAPILKEVGKKAALQVSMNDGLSFISSVITITTHCSDG<td>318</td></td>	KEPFSVEDTYLCPAPILKEVGKKAALQVSMNDGLSFISSVITITTHCSDG <td>318</td>	318

```
RESULT 3
A1R2_HUMAN STANDARD; PRT; 489 AA.
ID A1R2_HUMAN
AC P58335; O86U11; Q8NB13; Q96NC7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Anthrax toxin receptor 2 precursor (Capillary morphogenesis protein-2)
(GMG-2).
GN Name=ANTXR2; Synonyms=CMG2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
```

CC NCBI_TaID=9606;
RN [1]
RN SEQUENCE FROM N.A. (ISOFORM 2).
RX PubMed=21539596; PubMed=11693410;
RA Bell S.E., Mavila A., Salazar R., Bayless K.J., Kanagala S.,
RA Maxwell S.A., Davis G.E.;
RT "differential gene expression during capillary morphogenesis in 3D
RT collagen matrices: regulated expression of genes involved in basement
RT membrane matrix assembly, cell cycle progression, cellular
RT differentiation and G-protein signaling."
RL J. Cell Sci. 114:2755-2773(2001).
RN [2]
RN SEQUENCE FROM N.A. (ISOFORM 1), AND INTERACTION WITH ANTHRAX TOXIN.
RP TISSUE=Placenta;
RX MEDLINE=2608610; PubMed=12700348; DOI=10.1073/pnas.0431098100;
RA Scobie H.M., Rahney G.J.A., Bradley K.A., Young J.A.T.;
RT "Human capillary morphogenesis protein 2 functions as an anthrax toxin
RT receptor";
RL Proc. Natl. Acad. Sci. U.S.A. 100:5170-5174(2003).
RN [3]
RN SEQUENCE OF 78-489 FROM N.A. (ISOFORM 3), AND SEQUENCE FROM N.A.
RP (ISOFORM 4).
RC TISSUE=Synovial cell;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Makamatsu A., Hayashi K., Sato H., Nagai K., Kimura T., Makita H.,
RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Nagahata M., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA Omura Y., Abe K., Kamihara K., Katsuma N., Sato K., Tanikawa M.,
RA Yamazaki M., Nimomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe S., Hirooka S., Chiba Y.,
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yoshida M., Houta T.,
RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Togashi S., Komai F., Hara R., Takeuchi K., Arita M.,
RA Imose N., Mueselino K., Yuuki H., Oshima A., Sasaki N., Aotsuma S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shichata N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki M., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi T.,
RA Fujimori Y., Komiyama M., Taehiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiroo M., Okmoto Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikemori Y., Okamoto S.,
RA Ohtsuka R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togoashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuno Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs";
RUL Nat. Genet. 36:40-45(2004).
CC -1- FUNCTION: Cellular role is not yet known.
CC -1- SUBUNIT: Binds to the protective antigen (PA) of Bacillus
CC anthracis in a divalent cation-dependent manner, with the
CC following preference: calcium > manganese > magnesium > zinc.
CC Seems to bind to collagen type IV and laminin.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (isoforms 1 and 2).
CC Secreted (isoform 3). Isoform 1 is expressed at the cell surface
CC while isoform 2 is predominantly expressed within the endoplasmic
CC reticulum and not at the plasma membrane.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Name=1;
CC IsoId=P58335-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P58335-2; Sequence=VSP_008343;
CC Name=3;
CC IsoId=P58335-3; Sequence=VSP_008344, VSP_008345;
CC Note=No experimental confirmation available;

[illegible]

Db 275 NSMLCPAPILNKAGETLDVSVSPNGSKSVIGSLVITATECSNG 318

RESULT 4

Q6DPX2 PRELIMINARY; PRT; 487 AA.

AC Q6DPX2; 25-OCT-2004 (TReMBLrel. 28, Created)

DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)

DE 25-OCT-2004 (TReMBLrel. 28, Last annotation update)

DE Anthrax toxin receptor 2.

GN Name=Anthr2;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxId=10090;

[1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6; TISSUE=Eye;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Hsieh N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marubina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Uddin T.B., Toshlyuk S., Carninci P., Prange C.J., Bork S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W., Villalón D.K., Muzny D.C., Sodergren E.J., Lu X., Gibbs R.A., Paley J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S., Krzywinski M.I., Skalka U., Smalhe U., Schnerch A., Schein J.E., Jones S.J., Maira M.A.; "Generation and initial analysis of more than 15,000 full-length human RT and mouse cDNA sequences."

RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RL [2]

RN SEQUENCE FROM N.A.

RP STRAIN=C57BL/6; TISSUE=Eye;

RC Strauberg R.;

RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL: BC076595; AAH76595.1; --

DR GO: GO:0016021; C:Integral to membrane, IEA.

DR GO: GO:0004872; F:receptor activity; IEA.

DR InterPro: IPR008400; Anth_IG.

DR InterPro: IPR008399; Ant_C.

DR InterPro: IPR002035; VWF_A.

DR Pfam: PF05587; Anth_IG; 1.

DR Pfam: PF05586; Ant_C; 1.

DR SMART: SM00092; VWA; 1.

DR SMART: SM00337; VWA; 1.

DR PROSITE: PS50234; VWFA; 1.

KM Receptor.

SO SEQUENCE 487 AA; 53184 MW; 61A400D6BCBDE69 CRC64;

Query Match 50.5%; Score 770; DB 2; Length 487;

Best Local Similarity 52.8%; Pred. No. 1.8e-52;

Matches 150; Conservative 52; Mismatches 80; Indels 2; Gaps 1;

Qy 10 PACGSGDYLPIIDKSSGLAHNNEIYFPBQLAHKPTISPOLRNSPIYFSGRTTLMKLT 69

Db 37 PSCKKADLVFLVDKSSGVANNWLEIYNFVQLTERFVSPBMRSLFVFPSSQATIIILPT 96

Qy 70 EDREQIQGLELEIKVLPFGSDTYWGEGERASEQIYYENRGYRTASVITALTGELHED 129

Db 97 GDRKIKIGKLELDKAVKPVGETTYHESGKLANEQI--QNAQGLKASSIIILATGKLDGL 154

Qy 130 LFYSEREARNSRDGLAIYVCVGKDPNETQLARIADSKHVFVNDGFOALQGIHSIL 189

Db 155 VPSYAENBAKKSRSIGASVYCVGLDFEQALERIADSKQFPVKGCFQALKGIHSIL 214

Qy 190 KKSCEITLAAEPSTICGESFOVYVNRGNFPHANVDRVYCSFKINSVLTNEKPFVED 249

Db 215 AQSTETLELSPSSVCGEKFQVLTGRAVTSISHDSVLCCTFTANSTYTKSEKRVSIOP 274

Qy 250 TYLLCPAPILNKAGETLDVSVSPNGSKSVIGSLVITATECSNG 318

Db 275 NSMLCPAPILNKAGETLDVSVSPNGSKSVIGSLVITATECSNG 318

RESULT 5

Q8BVM2 PRELIMINARY; PRT; 641 AA.

AC Q8BVM2; 01-MAR-2003 (TReMBLrel. 23, Created)

DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)

DE 01-OCT-2003 (TReMBLrel. 25, Last annotation update)

DE Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:493430v1 product:hypothetical Prolin-rich region/von Willebrand factor type A domain containing protein, full insert sequence.

GN Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxId=10090;

[1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Testis;

RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;

RA Carninci P., Hayashizaki Y.;

RT "High-efficiency full-length cDNA cloning.";

RL Meth. Enzymol. 303:19-44(1999).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Testis;

RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;

RA RIKEN FANTOM Consortium;

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

RL [3]

RN SEQUENCE FROM N.A.

RP STRAIN=C57BL/6J; TISSUE=Testis;

RC The FANTOM Consortium.

RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";

RL Nature 420:563-573(2002).

RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Testis;

RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;

RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Komno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;

RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.";

RL Genome Res. 10:1617-1630(2000).

RN [5]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Testis;

RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;

RA Shibata K., Itoh M., Aizawa K., Nagaoke S., Sasaki N., Carninci P., Komno H., Akiyama U., Nishi K., Kitanai T., Tashiro H., Itoh M., Saito M., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A., Yamamoto R., Matsunoto H., Sakaguchi S., Ikegami T., Kashiwagi K., Fujisawa S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe M., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsumura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;

RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer.";

RL Genome Res. 10:1757-1771(2000).

RN [6]

```

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RA Adachi J., Alizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashimoto W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiroka T., Hirose T.,
RA Hori F., Imocani K., Ishii Y., Itoh M., Kagawa I., Kanakawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Saeki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takano-Akiba S., Tanaka Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK077206; BAC36683.1; -
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR008400; Anth_Ig.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF05587; Anth_Ig; 1.
DR Pfam; PF00092; VWA; 1.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS50234; VWF_A; 1.
DR Hypothetical protein.
SQ SEQUENCE 641 AA; 70415 MW; 199E300730BC85E3 CRC64;

Query Match 39.7%; Score 606; DB 2; Length 641;
Best Local Similarity 42.2%; Pred. No.2,3e-39;
Matches 119; Conservative 58; Mismatches 105; Indels 0; Gaps 0;

Qy 9 GPACGGFDLYFLIDKSGSVLHNNELIYFVBOIAHKFISPOLNRSFVSTRTGTTMKT 68
Db 68 GDCCQGIPLDLYLDKSGSVADNWHIYFAGLVKFKFTNNLISITVSTEARVLIPL 127
Qy 69 TEDREIQGJLELQKVLPGDDTVHMEGFERASRQIYENNGQRTASVIALDGEIHE 128
Db 128 TSDSEIKNSLLVLKSVIPQGLTHMQKGLRKANEIRKSTAGRIIVNVIITDGLLL 187
Qy 129 DLFFYSEERANRSRLGAIVYCVGVKDFNETOLARIADSKDHFVPVNDGFOALQIHSI 188
Db 188 KPYLDITMEAKKARMGALIVTVGVPMYSKOOLVNIADDPKRCVDSGFSALGVNDPL 247
Qy 189 LKKSCEILAAEPSTICAGESFQVVVRNGFRHARNVDRVLCSPKINDSVTLNKPESVE 248
Db 248 TSKSCTEILSVQPTVYVCAKDFYQVVISGHGINTSNMKGIVCRFKFSKVVDSPIDMN 307
Qy 249 DTVLLCPAPILKEVGMKALQVSMNDGISFTSSSVIITTHC 290
Db 308 EHSITCPGPKIKHGEDVSLQVSLNNGISFIGNKLIITSTNC 349

RESULT 6
Q96EC6 PRELIMINARY; PRT; 97 AA.
AC Q96EC6;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE ANTXR1 protein (Fragment).
GN Name=ANTXR1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast;
RX MEDLINE=2338257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold B.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheef C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh P.,

```

```

RA Diatchenko L., Matsushita K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Udén T.B., Tothiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abremson R.D., Mullaly S.J.,
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzyzanski M.I., Skalka U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.,
RT "generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast;
RA Strausberg R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC012475; AAH12475.1; -
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR008400; Anth_Ig.
DR Pfam; PF05587; Anth_Ig; 1.
DR NON TER 1
SQ SEQUENCE 97 AA; 10453 MW; 14F475F0B170E71A CRC64;

Query Match 25.4%; Score 388; DB 2; Length 97;
Best Local Similarity 95.1%; Pred. No.4,1e-23;
Matches 78; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 210 FOVVVRNGFRHARNVDRVLCSPKINDSVTLNKPESVEDTYLLCPAPILKEVGMKALQ 269
Db 1 FOVVVRNGFRHARNVDRVLCSPKINDSVTLNKPESVEDTYLLCPAPILKEVGMKALQ 60
Qy 270 VSMNDGISFTSSSVIITTHCS 291
Db 61 VSMNDGISFTSSSVIITTHCS 82

RESULT 7
ITAD RAT STANDARD; PRT; 1161 AA.
ID ITAD RAT
AC Q9QYB7;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Integrin alpha-D precursor.
GN Name=Itgad;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA O'Brien M.M., VanderVliet M., Klugman P.D., Dietrich G.,
RA Gallatin W.M.;
RT "Cloning of rat alpha D, a novel beta 2 integrin."
RL submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Integrin alpha-D/beta-2 is a receptor for ICAM3 and
CC VCAM1. May play a role in the atherosclerotic process such as
CC clearing lipoproteins from plaques and in phagocytosis of blood-
CC borne pathogens, particulate matter, and senescent erythrocytes
CC from the blood (By similarity).
CC -1- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-D
CC associates with beta-2 (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -1- DOMAIN: The integrin I-domain (inset) is a VWF domain. Integrins
CC with I-domains do not undergo protease cleavage.
CC -1- SIMILARITY: Belongs to the integrin alpha chain family.

```



```
CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
CC -1- SIMILARITY: Contains 1 VMPA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announce/
CC or send an email to license@ebi.ac.uk).
```

EMBL: AF021334; AAF21241.1; --
HBBP, P1215; 188Q.
InterPro: IPR000413; Integrin_alpha.
InterPro: IPR02035; VMP_A.
Pfam: PF01839; FG-GAP; 3.
Pfam: PF00357; Integrin_alpha; 1.
Pfam: PF00092; VMA; 1.
PRINTS: PR01185; INTEGRINA.
SMART: SM00453; VMPADOMAIN.
SMART: SM00327; VMA; 1.
PROSITE: PS00242; INTEGRIN_ALPHA; 1.
PROSITE: PS50234; VMPA; 1.
Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
Repeat; Signal; Transmembrane.
FT SIGNAL 1 19 Potential.
FT CHAIN 20 1161 Integrin alpha-D.
FT TRANSMEM 1101 1121 Extracellular (Potential).
FT DOMAIN 1122 1161 Potential.
FT REPEAT 34 87 Cytoplasmic (Potential).
FT REPEAT 88 7 FG-GAP 1.
FT DOMAIN 152 334 VMPA.
FT REPEAT 352 402 FG-GAP 3.
FT REPEAT 403 454 FG-GAP 4.
FT REPEAT 456 517 FG-GAP 5.
FT REPEAT 519 577 FG-GAP 6.
FT REPEAT 582 634 FG-GAP 7.
FT CA_BIND 467 475 Potential.
FT CA_BIND 531 539 Potential.
FT CA_BIND 594 602 Potential.
FT SITE 1126 1130 GPPKR motif.
FT DISULFID 69 76 By similarity.
FT DISULFID 108 126 By similarity.
FT DISULFID 656 711 By similarity.
FT DISULFID 769 775 By similarity.
FT DISULFID 845 860 By similarity.
FT DISULFID 993 1017 By similarity.
FT DISULFID 1022 1027 By similarity.
FT CARBOHYD 61 61 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 89 89 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 245 245 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 393 393 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 696 696 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 734 734 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 784 784 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 907 907 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 936 936 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1045 1045 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 1161 AA; 126600 MW; 2258491984A705E CRC64;

Query Match 10.5%; Score 159.5; DB 1; Length 1161;
Best Local Similarity 27.9%; Pred. NO. 0.00098;
Matches 63; Conservative 41; Mismatches 93; Indels 29; Gaps 11;

Qy 17 DLVEFLDKSGSV-LHHNWEIYFVEQLAKRISPOLRMSPIVPSFGTTLAKLKE----- 70
Db 152 DIATLIDSGSSINGRDFAQMDVFKALMGSEFASITLSLMQYINILKTHFTFEFNKIL 211
Qy 71 DREQIRQGLEBLQKLPQGGDTYMEGFERASQIYENRQGYRTA-SVIALTLTGSELHED 129
Db 212 DQSLVDPPIVQLQ-----GLTYTATGIRTVMEELFHSKNGSKSAKKILVITDQGYRD 266

```
Qy 130 LFPYSE--REANRSRDGAIVYCVGYVD-FNE-TQLARI-----ADSKDHVPVNDGFOA 180  
Db 267 PLRSDVLPADNKA---GIIRYALGVDAFQEPALKEMLNTIGSAPQDHFYKGN-FAA 322  
Qy 181 LQGIHSLKSKSCREIILAAEPSTICAGESFQVYVNGGFRRARVD 226  
Db 323 LRSIQRLQERK----IPAIKBTQSRSSSFQHEMSQGFSSALTSND 364
```

RESULT 8
ITAD HUMAN
ID ITAD_HUMAN STANDARD; PRT; 1162 AA.
AC Q13349; Q15575; Q15576;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Integrin alpha-D precursor (Leukointegrin alpha D) (CD11d) (ADB2).
GN Name=ITGAD;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=9611356; PubMed=8777714; DOI=10.1016/1074-7613(95)90058-6;
RA Van der Vlieten M., Le Trong H., Wood C.L., Moore P.F., St John T.,
RT Staunton D.B., Gallatin W.M.;
RL "A novel leukointegrin, alpha d beta 2, binds preferentially to ICAM-3."
RN [2]
RP IMMUNITY 3:683-690(1995).
RX MEDLINE=20187620; PubMed=10722744; DOI=10.1074/jbc.275.12.8959;
RA Noci J.D., Johnson A.K., Dillon J.D.;
RT "Structural and functional characterization of the leukocyte integrin gene CD11d. Essential role of Spi and Sp3.";
RL J. Biol. Chem. 275:8959-8969(2000).
RN [3]
RP SEQUENCE OF 467-537; 571-602; 633-666; 788-834 AND 910-1125 FROM N.A.
RX MEDLINE=96257236; PubMed=8666289; DOI=10.1016/0378-1119(95)00869-1;
RA Wong D.A., Davis E.M., Lebeau M., Springer T.A.;
RT "Cloning and chromosomal localization of a novel gene encoding a human beta 2-integrin alpha subunit.";
RL Gene 171:291-294(1996).
RN [4]
RP INTERACTION WITH VCAM1.
RX MEDLINE=99059842; PubMed=9841932;
RA Grayson M.H., Van der Vlieten M., Sterbinsky S.A., Michael Gallatin W.,
RT Hoffman P.A., Staunton D.E., Bochner B.S.;
RL "alpha2beta2 integrin is expressed on human eosinophils and functions as an alternative ligand for vascular cell adhesion molecule 1 (VCAM-1)."
RN [5]
RP J. Exp. Med. 188:2187-2191(1998).
RX INTERACTION WITH VCAM1.
RX MEDLINE=99370002; PubMed=10438935;
RA Van der Vlieten M., Crowe D.T., Hoekstra D., Vazeux R., Hoffman P.A.,
RT Grayson M.H., Bochner B.S., Gallatin W.M., Staunton D.E.;
RL "The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a binding interface between I domain and VCAM-1.";
RN J. Immunol. 163:1984-1990(1999).
RX -1- FUNCTION: Integrin alpha-D/beta-2 is a receptor for ICAM3 and VCAM1. May play a role in the atherosclerotic process such as clearing lipoproteins from plaques and in phagocytosis of blood-borne pathogens, particulate matter, and senescent erythrocytes from the blood.
CC -1- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-D associates with beta-2.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: Expressed moderately on myelomonocytic cell lines and subsets of peripheral blood leukocytes and strongly on

CC tissue-specialized cells, including macrophages foam cells within
CC atherosclerotic plaques, and on splenic red pulp macrophages.
CC -1- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC with I-domains do not undergo protease cleavage.
CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
CC -1- SIMILARITY: Contains 1 VWFA domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.ebi.ac.uk/announcements/](http://www.ebi.ac.uk/announcements)
CC or send an email to license@ebi.ac.uk).

CC -----
CC EMBL: U37028; AAB38547.1; -.
CC EMBL: U40274; AAB60634.1; -.
CC EMBL: U40275; AAB60635.1; -.
CC EMBL: U40276; AAB60636.1; -.
CC EMBL: U40277; AAB60637.1; -.
CC EMBL: U40279; AAB60638.1; -.
CC EMBL: U40278; AAB60638.1; JOINED.
CC EMBL: AF187881; AAF62875.1; -.
CC HSSP: P11215; 1BHQ.
CC Genew: HGNC:6146; ITGAD.
CC MIM: 602453; -.
CC GO: GO:0008305; C:integrin complex; TAS.
CC GO: GO:0016337; P:cell-cell adhesion; NAS.
CC GO: GO:0007160; P:cell-matrix adhesion; NAS.
CC GO: GO:0006955; P:immune response; NAS.
CC InterPro: IPR000413; Integrin_alpha.
CC InterPro: IPR02035; VWF_A.
CC Pfam: PF01839; FG-GAP; 3_
CC Pfam: PF00357; Integrin_alpha; 1.
CC Pfam: PF00092; VWA; 1.
CC PRINTS: PRO1185; INTEGRINA.
DR PRINTS: PRO0453; VWFADOMAIN.
DR SMART: SM00191; Int_alpha; 5.
DR SMART: SM00327; VMA; 1.
DR PROSITE: PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE: PS50234; VWFA; 1.
KW Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
KW Repeat; Signal; Transmembrane.
FT CHAIN 1 17 Potential.
FT DOMAIN 18 1162 Integrin_alpha-D.
FT TRANSMEM 1101 1121 Extracellular (Potential).
FT DOMAIN 1122 1162 Potential.
FT REPEAT 32 85 Cytoplasmic (Potential).
FT REPEAT 86 ? FG-GAP 1.
FT DOMAIN 150 332 FG-GAP 2.
FT REPEAT 350 400 VWFA.
FT REPEAT 401 452 FG-GAP 3.
FT REPEAT 454 516 FG-GAP 4.
FT REPEAT 518 576 FG-GAP 5.
FT REPEAT 581 633 FG-GAP 6.
FT REPEAT 645 691 FG-GAP 7.
FT CA_BIND 465 473 Potential.
FT CA_BIND 530 538 Potential.
FT CA_BIND 593 601 Potential.
FT SITE 1127 1131 GFFKR motif.
FT DISULFID 67 74 By similarity.
FT DISULFID 106 124 By similarity.
FT DISULFID 655 710 By similarity.
FT DISULFID 769 775 By similarity.
FT DISULFID 846 861 By similarity.
FT DISULFID 994 1018 By similarity.
FT DISULFID 1023 1028 By similarity.
FT CARBOHYD 59 59 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 87 87 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 99 99 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 391 391 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 691 691 N-linked (GlcNAc...) (Potential).

FT CARBOHYD 733 733 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 873 873 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 957 957 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1046 1046 N-linked (GlcNAc...) (Potential).
FT CONFLICT 500 500 Missing (in Ref. 2).
FT CONFLICT 515 518 GHW -> ATP (in Ref. 2).
FT CONFLICT 825 825 L -> V (in Ref. 2).
FT CONFLICT 984 984 V -> A (in Ref. 2).
SQ SEQUENCE 1162 AA; 12685 MW; F296A1A35455D77D CRC64;
Query Match 10.2%; Score 156; DB 1; Length 1162;
Best Local Similarity 24.3%; Pred. No. 0.0019;
Matches 67; Conservative 50; Mismatches 119; Indels 40; Gaps 11;
Qy 7 DGGPAC-YGGFDLYFLIDKSGSV-LHNMWYFYVQLAHKFTSPQLRMSFYFSTRTGT 64
Db 139 DATPECPHQMIDIVFLIDSGSIDQNDPFQMKGFVQAVMQFGDTLTFALMQYS-----N 194
Qy 65 LMKLTEDREQIRGLBELQVLP-----GGDTYHBEFEPASBQIYVENQGYRTA-SVIT 119
Db 195 LKIHFTFTQFRTSPSQSLVDEIVQLKGLFTATGILTVVTGLFHHKNGARRSAKKIL 254
Qy 120 ALTDGELHEDLPYSSREANRSRLGAIYCVGVDFNETQLAR-----IDSKDHF 172
Db 255 VITDGGKXKDPLEYSD-VLPQAEKAGIIRYAIVGHAFGQPTKROELNTSSAPPDHVF 313
Qy 173 PVNDGFQALGIIHSLIKSCIEILAEPSITGAGESFQVVRNGFRHARNDVRLC-- 230
Db 314 KV-DNFALGSIKQIQEK-----IYAVEGTQSASASSFQHEMQEGHFTLMDGLFLGA 368
Qy 231 --SFKINDS-----VTLNEKPSVEPTYL 252
Db 369 VGSFWSWGAFLYPPNMSPTFINMSQENVMDRSTYL 404
RESULT 9
Q9BP08 PRELIMINARY; PRT; 1332 AA.
AC Q9BP08
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Integrin alpha H1 precursor.
GN Name=HITGAL;
OS Halocynthia roretzi (Sea squirt).
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Stolidobranchia; Pyrosidae; Halocynthia.
OX NCBI_TaxID=7729;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hemocyte;
RA MEDLINE=21103187; PubMed=11160215;
RX Miyazawa S., Azumi K., Nonaka M.;
RT "Cloning and characterization of integrin alpha subunits from the
RT solitary ascidian, Halocynthia roretzi.";
RL J. Immunol. 166:1710-1715(2001).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (by similarity).
CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
DR EMBL: AB048261; BAB21479.1; -.
DR HSSP: P11215; 1BHQ.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0008305; C:integrin complex; IEA.
DR GO: GO:0005515; F:protein binding; IEA.
DR GO: GO:0007160; P:cell-matrix adhesion; IEA.
DR GO: GO:0007229; P:integrin-mediated signaling pathway; IEA.
DR InterPro: IPR000413; Integrin_alpha.
DR InterPro: IPR02035; VWF_A.
DR Pfam: PF00092; VWA; 1.
DR PRINTS: PRO1185; INTEGRINA.
DR PRINTS: PRO0453; VWFADOMAIN.
DR SMART: SM00191; Int_alpha; 5.
DR SMART: SM00327; VMA; 1.
DR PROSITE: PS50234; VWFA; 1.

KW Cell adhesion; Integrin; Signal; Transmembrane.
FT SIGNAL 1 30 Potential.
FT CHAIN 31 1332 Integrin alpha H1.
SQ SEQUENCE 1332 AA; 145851 MW; 0D9108D2B05CFEAE CRC64;
Query Match 10.1%; Score 153.5; DB 2; Length 1332;
Best Local Similarity 24.1%; Pred. No. 0.0035;
Matches 63; Conservative 42; Mismatches 81; Indels 75; Gaps 10;
QY 2 GGRREDG-----PACYGFDLYFLDKSGSVLHHMNEIYFVEQLAHKFISPOLRMSFI 56
DB 188 GNMREGSRNTECP--SGVDLFLVLDGSGSVGRKFDKVKMVKIT----- 232
QY 57 VFSTRGTTLMKLTEDREIQRLGELQKVLPG-----GDTYMHGFERAS 101
DB 233 -----AKLDIKELIVRGVQVSHYVEGKSINKQYITTEISIGFKLIDFENAV 283
QY 102 EQLIYENRGYRT-----ASVIALTDGELHEDLFFISEKARS 141
DB 284 DRI--QLQGYTTYGRALQKVIKRDPAIYGNKQVLLLTLDGQAKDKILP--NANRL 338
QY 142 RDLGAIYVCVGVDFNETOLARIA--DSKDHVPVNDGFOALQGIHSILKSCIEILA 198
DB 339 RAKGATFVAVGVDISELKLIAAGTSTDRVFTVTD--FGELDSIVSLQTEIQSFVLE 397
QY 199 AEPSTICAGESFQVVRNGNF 219
DB 398 CGKSAKTAG--YEMHFGENG 416
RESULT 10
Q8NFW1 PRELIMINARY; PRT; 1626 AA.
AC Q8NFW1,
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Alpha 1 type XXII collagen.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cartilage;
RA Koch M., Jin M., Ashworth T., Burgeon R.E.;
RL Submitted (AUG-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL, AF406780; AA03620.1; -.
DR HSSP, P18614; IIMP.
DR GeneW, HGNC:22989; COL22A1.
DR GO, GO:0005737; C:cytoplasm; IEA.
DR GO, GO:0005198; F:structural molecule activity; IEA.
DR GO, GO:0007155; P:cell adhesion; IEA.
DR GO, GO:0006817; P:positive transport; IEA.
DR InterPro, IPR008161; C1g helix.
DR InterPro, IPR008160; C1g helix.
DR InterPro, IPR008985; Cons. like _lec_g1.
DR InterPro, IPR003129; TSP_N.
DR InterPro, IPR002035; VWF_A.
DR Pfam, PF01391; Collagen_16.
DR Pfam, PF00092; VMA; 1.
DR PRINTS, PR00453; VWFADOMAIN.
DR PRODOM, PD000007; C1g helix; 7.
DR SMART, SM00210; TSPN_1.
DR SMART, SM00327; VMA; 1.
DR PROSITE, PS50234; VMA; 1.
KW Collagen.
SQ SEQUENCE 1626 AA; 16115 MW; 34C68E3CFD467407 CRC64;
Query Match 9.7%; Score 148.5; DB 2; Length 1626;
Best Local Similarity 25.4%; Pred. No. 0.011;
Matches 65; Conservative 45; Mismatches 105; Indels 41; Gaps 14;

QY 8 GGPACYG-----GPDLYFLDKSGSV-LHHMNEIYFVEQLAHKF-ISP-QLRMSFI 56
DB 21 GGGCGQAOGRACCKSVHYDLVFLDPTSSSVGKEDEKRVQWVAVNVDTEVGPDRTRVGVV 80
QY 57 VFSTRGTTLMK--LTEDREIQRLGELQKVLPGDPTYMHGFERASQIYENRG-- 111
DB 81 RYSDRPTTAFELGIFGSOEEVKAARRL--AYHGANTTGALRYITARSFSPHAGGPR 138
QY 112 ---YRTASVIALTDGELHEDLFFYSERARSRLGAIYVCVGVDFNETOLARIAD-- 166
DB 139 DRAK--QVALLTDGR--SGLVVDAAAAARA---GIRFVAGELAKLELEISAP 192
QY 167 SKDHVPVNDGFOALQGIHSILKSCIEILAEPSTICAGESFQVVRNGFPRHARNVD 226
DB 193 KSAHVFHVSD--FNALDKIRGLRRRLCENVLC--PS-----VRVEGDRFHTNGT 240
QY 227 RVLCSPKINDSVTINE 242
DB 241 KEITGFDLMDLFVKE 256
RESULT 11
Q8T6U5 PRELIMINARY; PRT; 441 AA.
AC Q8T6U5,
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Proximal thread matrix protein 1 variant a.
OS Mytilus edulis (Blue mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_Taxid=6550;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22313036; PubMed=12425661; DOI=10.1021/bm0255903;
RA Sun C., Lucas J.M., Waite J.H.;
RT "Collagen-binding matrix proteins from elastomeric extraorganic
RT bivalve fibers."
RL Biomacromolecules 3:1240-1248(2002).
DR EMBL, AF414454; AA483537.1; -.
DR GO, GO:0005198; F:structural molecule activity; IEA.
DR InterPro, IPR02035; VWF_A.
DR Pfam, PF00092; VMA; 2.
DR PRINTS, PR00453; VWFADOMAIN.
DR SMART, SM00327; VMA; 2.
DR PROSITE, PS50234; VMA; 2.
KW Matrix protein.
SQ SEQUENCE 441 AA; 47543 MW; 881D8BD36B891D2B CRC64;
Query Match 9.7%; Score 148; DB 2; Length 441;
Best Local Similarity 25.5%; Pred. No. 0.0025;
Matches 51; Conservative 46; Mismatches 77; Indels 26; Gaps 12;
QY 12 CYGFDLYFLDKSGSV-----LHHMNEIYFVEQLAHKF--ISP-LRMSFIVSTRGTT 64
DB 236 CAGNADIAFPDASSSINANNPNYGLMKDFMCDIVRFKGTGDDGQFPAVTFADADATK 295
QY 65 LMKLTE--DREQIRQGLGELQKVLPG--GDTYMHGFERASQIYENRG--YRTASV 117
DB 296 QFGIKDYSKAEIKGAID---KVPSTIIGQTALDGIENARLEV--FNNRGGGGRREVQKV 351
QY 118 IIALTDGEL--HEDLFFYSERARSRLGAIYVCVGV--KDFNETOLARIADSDHYEPV 174
DB 352 VILLTDQNNNGHS---PEHESLNRKEGVVAIVAGVGFLKSELINIASSEIVF--T 406
QY 175 NDGFQALQGIHSILKSCI 194
DB 407 TSSFNKLSTIMENVVYKACM 426
RESULT 12
Q8T5C3

RT assessment." ;
 RL Structure 6:923-935(1998).
 RN [12]
 RP 3D-STRUCTURE MODELING OF 17-616.
 RX MEDLINE=98226734; PubMed=9560195; DOI=10.1073/pnas.95.9.4870;
 RA Oxyg C., Springer T.A.;
 RT "Experimental support for a beta-propeller domain in integrin alpha-subunits and a calcium binding site on its lower surface." ;
 RL Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875(1998).
 CC -1 FUNCTION: Integrin alpha-M/beta-2 is implicated in various adhesive interactions of monocytes, macrophages and granulocytes as well as in mediating the uptake of complement-coated particles. It is identical with CR-3, the receptor for the IC3b fragment of the third complement component. It probably recognizes the R-G-D peptide in C3b. Integrin alpha-M/beta-2 is also a receptor for fibrinogen, factor X and ICAM1. It recognizes P1 and P2 peptides of fibrinogen gamma chain.
 CC -1 SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-M associates with beta-2.
 CC -1 SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1 TISSUE SPECIFICITY: Predominantly expressed in monocytes and granulocytes.
 CC -1 DOMAIN: The integrin I-domain (insert) is a WFPA domain. Integrins with I-domains do not undergo protease cleavage.
 CC -1 SIMILARITY: Belongs to the integrin alpha chain family.
 CC -1 SIMILARITY: Contains 7 FG-GAP repeats.
 CC -1 SIMILARITY: Contains 1 WFPA domain.
 CC -1 DATABASE: NAME=PROV; NOTE=CD guide CD11b entry; WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd11b.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.ebi.ac.uk/announcements or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL/ J03925; AAA59544.1; -;
 DR EMBL/ M18044; AAA59491.1; -;
 DR EMBL/ J04145; AAA59903.1; -;
 DR EMBL/ S52227; AAB24821.1; -;
 DR EMBL/ S52152; AAB24821.1; JOINED.
 DR EMBL/ S52153; AAB24821.1; JOINED.
 DR EMBL/ S52154; AAB24821.1; JOINED.
 DR EMBL/ S52155; AAB24821.1; JOINED.
 DR EMBL/ S52157; AAB24821.1; JOINED.
 DR EMBL/ S52159; AAB24821.1; JOINED.
 DR EMBL/ S52161; AAB24821.1; JOINED.
 DR EMBL/ S52164; AAB24821.1; JOINED.
 DR EMBL/ S52165; AAB24821.1; JOINED.
 DR EMBL/ S52167; AAB24821.1; JOINED.
 DR EMBL/ S52169; AAB24821.1; JOINED.
 DR EMBL/ S52170; AAB24821.1; JOINED.
 DR EMBL/ S52173; AAB24821.1; JOINED.
 DR EMBL/ S52174; AAB24821.1; JOINED.
 DR EMBL/ S52180; AAB24821.1; JOINED.
 DR EMBL/ S52181; AAB24821.1; JOINED.
 DR EMBL/ S52184; AAB24821.1; JOINED.
 DR EMBL/ S52189; AAB24821.1; JOINED.
 DR EMBL/ S52191; AAB24821.1; JOINED.
 DR EMBL/ S52192; AAB24821.1; JOINED.
 DR EMBL/ S52203; AAB24821.1; JOINED.
 DR EMBL/ S52212; AAB24821.1; JOINED.
 DR EMBL/ S52213; AAB24821.1; JOINED.
 DR EMBL/ S52216; AAB24821.1; JOINED.
 DR EMBL/ S52219; AAB24821.1; JOINED.
 DR EMBL/ S52220; AAB24821.1; JOINED.
 DR EMBL/ S52221; AAB24821.1; JOINED.
 DR EMBL/ S52222; AAB24821.1; JOINED.
 DR EMBL/ S52226; AAB24821.1; JOINED.
 DR EMBL/ M76724; AAA58410.1; -;
 DR EMBL/ M84477; AAA51960.1; -;

DR PIR; A31108; RWHUB.
 DR PDB; 1ABX; Model; @=17-1152.
 DR PDB; 1BHO; X-ray; 1/2-
 DR PDB; 1BHQ; X-ray; 1/2-
 DR PDB; 1IDN; X-ray; 1/2-
 DR PDB; 1IDO; X-ray; @=140-331.
 DR PDB; 1JLM; X-ray; @=143-334.
 DR PDB; 1MLU; X-ray; A=137-331.
 DR PDB; 1MF7; X-ray; A=144-337.
 DR PDB; 1N9Z; X-ray; A=140-335.
 DR PDB; 1NA5; X-ray; A=144-345.
 DR Genew; HGNC:6143; ITGAM.
 DR MIM; 120980; -;
 DR GO; GO:0008305; C:integrin complex; TAS.
 DR GO; GO:0007155; P:cell adhesion; TAS.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00357; Integrin_alpha; 1.
 DR Pfam; PF00092; VWA; 1.
 DR PRINTS; PR01185; INTEGRINA.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00191; Int_alpha; 5.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS50234; WFPA; 1.
 DR 3D-structure; Calcium; Cell adhesion; Direct protein sequencing;
 KW Glycoprotein; Integrin; Magnesium; Receptor; Repeat; Signal;
 KW Transmembrane.
 FT SIGNAL 1 16
 FT CHAIN 17 1152 Integrin alpha-M.
 Query 17 DLYEILDKSGSVL-HHNNIYFYVEQLAHKFIQSPRLMSFVSTRTGTLTKLTD----- 71
 Db 150 DIAFLIDGSGSIIPHDFRMRKEFVST-----VMEQLKSKTLFS-----LMQYSEBRRIH 199
 Qy 72 -----REQIROLGEELOKVLPGCDTVMHEGERASEQIYYENRGYRTA-SVITALT 122
 Db 200 FTPEKPNPNPNSPLVPIQLL--GRTHATATGIRKVRRELFNTNGARKNAFKILVIT 257
 Qy 123 DGEIHELDLPYSE--REANSRDLGATVCGVVDENETOLAR-----IADS--KHVFP 173
 Db 258 DGEKFGDPLGVEDVTEPADRB--GVIRVYIGVDAPRSEKSRQELNTATSKPRDHFQ 314
 Qy 174 VNDGFQALOGIILHSILKSCIEILAAEPSTICAGESFQVAVRGNGFRA 222
 Db 315 VNN-FEALKTITQNGLRK-----IPAIETQTGSSSSFEHMSQEGFSA 358
 RESULT 14
 OBTSC2 PRELIMINARY; PRT; 453 AA.
 AC OBTSC2;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Proximal thread matrix protein 1.
 OS Mytilus galloprovincialis (Mediterranean mussel).
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
 OC Mytiloidea; Mytilidae; Mytilus.
 OC NCBI_TaxID=29158;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22313036; PubMed=12425661; DOI=10.1021/bm0255903;
 RA Sun C., Lucas J.M., Waite J.H.;
 RT "Collagen-binding matrix proteins from elastomeric extraorganic
 RT byssal fibers." ;
 RL Biomacromolecules 3:1240-1248(2002).
 DR EMBL; AY053391; AAL17974.1; -;

DR HSSP; P20701; IMCN.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR02035; VWF_A.
DR Pfam; PF00092; VWA; 2.
DR PRINTS; PRO0453; VWFADOMAIN.
DR SMART; SM00327; VWA; 2.
DR PROSITE; PS50234; VWF_A; 2.
DR Matrix protein.
KM SEQUENCE 453 AA; 48784 MW; D60497F5C0C51E6D CRC64;
SQ
Query Match 9.5%; Score 145; DB 2; Length 453;
Best Local Similarity 25.5%; Pred. No. 0.0044;
Matches 51; Conservative 44; Mismatches 79; Indels 26; Gaps 12;
QY 12 CYGGEFDYFIDKSGSV-----LHMNEIYFVEQLANHF--ISFQ-LRMSFVSTGTT 64
DB 248 CAGADIAFVDFADASSINANNPNVQLMKNFMKIIVDRFKTGPDGQFAVTFADATK 307
QY 65 LMKLTE--DREQIRQGLEELQKVLPG--GDTYMHGEFPERASEQIYENRQG--YRTASV 117
DB 308 QFGLEKDYSSKADIKGAID---KVSPSIIIGQTAIGDGLNARLEV-FPRNNGGHEEVQKV 363
QY 118 IIALTDGEL--HEDLFFSREARSRDLGAIYVCVQ-KQNFQTQARLSDQHPV 174
DB 364 VILLTDGQNGHKS---PEHESLLRKEGVVAIGVGTGFLKSLINIASSEYVF-T 418
QY 175 NDGFQALQGIHSILKKSCI 194
DB 419 TSSFDKLSKIMEDVYKLACM 438
RESULT 15
ITAX HUMAN STANDARD; PRT; 1163 AA.
ID _TAX_HUMAN
AC P20702; OSIVA6; 17, Created
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Integrin alpha-X precursor (Leukocyte adhesion glycoprotein p150,95
DE alpha chain) (Leukocyte adhesion receptor p150,95) (CD11c) (Lew M5).
GN Name=ITGAX; Synonyms=CD11C;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9816645; PubMed=3327687;
RA Cordi A.L., Miller L.J., O'Connor K., Larson R.S., Springer T.A.;
RT "cDNA cloning and complete primary structure of the alpha subunit of a
RT leukocyte adhesion glycoprotein, p150,95.";
RL EMBO J. 6:4023-4028(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90153906; PubMed=2303426;
RA Cordi A.L., Garcia-Aguilar J., Springer T.A.;
RT "Genomic structure of an integrin alpha subunit, the leukocyte p150,95
RT molecule.";
RL J. Biol. Chem. 265:2782-2788(1990).
RN [3]
RP ERRATUM.
RA Cordi A.L., Garcia-Aguilar J., Springer T.A.;
RL J. Biol. Chem. 265:12750-12751(1990).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Straube R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datschenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,

RA Brownstein M.J., Ueda T.B., Toshiyuki S., Garnick P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Pahey J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP SEQUENCE OF 20-43.
RX MEDLINE=87167596; PubMed=3549901;
RA Miller L.J., Wiebe M., Springer T.A.;
RT "Purification and alpha subunit N-terminal sequences of human Mac-1
RT and p150,95 leukocyte adhesion proteins.";
RL J. Immunol. 138:2381-2383(1987).
CC -1- FUNCTION: Integrin alpha-X/beta-2 is a receptor for fibrinogen. It
CC recognizes the sequence G-P-R in fibrinogen. It mediates cell-cell
CC interaction during inflammatory and chemotaxis.
CC important in monocyte adhesion and chemotaxis.
CC -1- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-X
CC associates with beta-2.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: Predominantly expressed in monocytes and
CC granulocytes.
CC -1- DOMAIN: The integrin I-domain (insert) is a VWA domain. Integrins
CC with I-domains do not undergo protease cleavage.
CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
CC -1- SIMILARITY: Contains 1 VWA domain.
CC -1- DATABASE: NAME=PRO; NOTE=CD guide CD11c entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cdlc.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; M81695; AA59180.1; -;
DR EMBL; M29165; -; NOT ANNOTATED CDS.
DR EMBL; M29487; AA51620.1; ALT_SEQ.
DR EMBL; M29482; AA51620.1; JOINED.
DR EMBL; M29483; AA51620.1; JOINED.
DR EMBL; M29484; AA51620.1; JOINED.
DR EMBL; M29485; AA51620.1; JOINED.
DR EMBL; M29486; AA51620.1; JOINED.
DR EMBL; BC038237; AA38237.1; -;
DR PIR; A36584; RWHUIC.
DR PDB; 1N3Y; X-ray; A=141-338.
DR GeneW; HGNC:6152; ITGAX.
DR MIM; 151510; -;
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0004872; F:receptor activity; TAS.
DR GO; GO:0007155; P:cell adhesion; TAS.
DR GO; GO:0009887; P:organogenesis; TAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3_
DR Pfam; PF00357; Integrin_alpha; 1.
DR Pfam; PF00092; VWA; 1.
DR PRINTS; PRO1185; INTEGRINA.
DR PRINTS; PRO0453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.

DR PROSITE, PS50234, VMPA, 1.
 KW 3D-structure; Calcium; Cell adhesion; Direct protein sequencing;
 KW Glycoprotein; Integrin; Magnesium; Polymorphism; Receptor; Repeat;
 KW Signal; Transmembrane.
 FT SIGNAL 1 19
 FT CHAIN 20 1163
 FT DOMAIN 20 1107
 FT TRANSMEM 1108 1128
 FT DOMAIN 1129 1163
 FT REPEAT 34 87
 FT REPEAT 88 ?
 FT DOMAIN 165 351
 FT REPEAT ? 401
 FT REPEAT 402 453
 FT REPEAT 455 517
 FT REPEAT 518 576
 FT REPEAT 581 633
 FT CA_BIND 466 474
 FT CA_BIND 530 538
 FT CA_BIND 593 601
 FT SITE 1131 1135
 FT DISULFID 69 76
 FT DISULFID 108 126
 FT DISULFID 655 712
 FT DISULFID 771 777
 FT DISULFID 848 863
 FT DISULFID 998 1022
 FT DISULFID 1027 1032
 FT CARBOHYD 61 61
 FT CARBOHYD 89 89
 FT CARBOHYD 392 392
 FT CARBOHYD 697 697
 FT CARBOHYD 735 735
 FT CARBOHYD 899 899
 FT CARBOHYD 939 939
 FT CARBOHYD 1050 1050
 FT VARIANT 48 48
 FT CONFLICT 209 209
 FT CONFLICT 251 251
 FT CONFLICT 469 469
 FT CONFLICT 490 490
 FT CONFLICT 547 547
 FT CONFLICT 756 756
 FT CONFLICT 819 819
 FT CONFLICT 1163 1163
 FT STRAND 150 157
 FT TURN 160 161
 FT HELIX 164 178
 FT TURN 179 180
 FT TURN 183 185
 FT STRAND 186 193
 FT STRAND 197 201
 FT HELIX 203 208
 FT TURN 212 216
 FT STRAND 217 218
 FT HELIX 226 226
 FT STRAND 228 236
 FT TURN 237 240
 FT HELIX 242 244
 FT TURN 245 245
 FT TURN 248 249
 FT STRAND 251 258
 FT STRAND 263 263
 FT HELIX 269 278
 FT TURN 279 280
 FT STRAND 282 288
 FT HELIX 290 293
 FT TURN 296 297
 FT HELIX 298 304
 FT HELIX 310 312
 FT STRAND 313 316
 FT HELIX 319 325

Integrin alpha-X.
 Extracellular (Potential).
 Potential.
 Cytoplasmic (Potential).
 FG-GAP 1.
 FG-GAP 2.
 VMPA.
 FG-GAP 3.
 FG-GAP 4.
 FG-GAP 5.
 FG-GAP 6.
 FG-GAP 7.
 Potential.
 Potential.
 Potential.
 GFPKR motif.
 By similarity.
 By similarity.
 By similarity.
 By similarity.
 By similarity.
 By similarity.
 By similarity.
 N-linked (GlcNAc. . .) (Potential).
 N-linked (GlcNAc. . .) (Potential).
 N-linked (GlcNAc. . .) (Potential).
 N-linked (GlcNAc. . .) (Potential).
 N-linked (GlcNAc. . .) (Potential).
 N-linked (GlcNAc. . .) (Potential).
 N-linked (GlcNAc. . .) (Potential).
 N-linked (GlcNAc. . .) (Potential).
 N-linked (GlcNAc. . .) (Potential).
 W -> R (in dbSNP:11574633).
 /FTid=VAR_018672.
 T -> S (in Ref. 4).
 T -> A (in Ref. 4).
 T -> S (in Ref. 4).
 G -> A (in Ref. 2).
 E -> K (in Ref. 4).
 D -> L (in Ref. 1).
 I -> V (in Ref. 4).
 SEK -> TPHPQDNV (in Ref. 4).

FT HELIX 326 334
 FT TURN 335 335
 SQ SEQUENCE 1163 AA; 127886 MW; 83658A13B5C5DE8F CRC64;
 Query Match 9.4%; Score 143.5; DB 1; Length 1163;
 Best local similarity 24.8%; Pred. No. 0.018; Indels 35; Gaps 10;
 Matches 55; Conservative 43; Mismatches 89;
 QY 17 DLYFILDKSGSV-LHHNNEIYFVEQLAHKFIQPLMSFVIFSTGTTLTKLTEDRE-- 73
 DB 151 DIVLIDSGSGSISRNPFATMNFVRAVISQFRSTQPSLMQFENKQTHTFEERFRTS 210
 QY 74 ---QIRQGLEFLQVVLREGDTYMEGFERSAQIYENRQGYRTAS-VTIALTQGEIHED 129
 DB 211 NPLSLASVHQLQ-----GFTYTATAIQNVVHRLPHASYGARRDATKILIVITDGKKEGD 265
 QY 130 LFFYSEREARNSRLGALVYCVGV-----KDNENQGLARIAD--SKDHVEPVVDG 177
 DB 266 SLDYKD-VTPMADAGIIRYALGVGLAPQENNSKELND-----TASKPSQEHIFKVED- 318
 QY 178 FOALQGIHSHILKSCIEILAEPSITCAGESFOVVRGNGF 219
 DB 319 FDALKDQIONLQKEK---IFAIEGTETTSSTSSFELEMAQEGF 356

Search completed: June 13, 2005, 20:03:14
 Job time : 99.3587 secs

THIS PAGE BLANK (USPTO)


```
XX The invention relates to human nucleic acids (AA157798-AA161369) and the
CC encoded polypeptides (AA038642-AA042213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 297 AA;
Query Match 100.0%; Score 970; DB 4; Length 297;
Best Local Similarity 100.0%; Pred. No. 1.6e-99;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGFPLVFLDKSGSVLHHNNEIYFVPEQLAHKFTSPQLRMSFIYFSTRGTTLMKLTEDRE 60
DB 41 GGFPLVFLDKSGSVLHHNNEIYFVPEQLAHKFTSPQLRMSFIYFSTRGTTLMKLTEDRE 100
QY 61 QIRGSELEOKVLPGGDTYHMEGFERRASEQIYYENRQGYRTASVITALTDEGLHEDLFFY 120
DB 101 QIRGSELEOKVLPGGDTYHMEGFERRASEQIYYENRQGYRTASVITALTDEGLHEDLFFY 160
QY 121 SEREANSRDLGAIYVCVGVDFNETQLARIADSKDHVFVNDGFQALGGIHSILKKSC 180
DB 161 SEREANSRDLGAIYVCVGVDFNETQLARIADSKDHVFVNDGFQALGGIHSILKKSC 220
QY 181 IEIILAE 187
DB 221 IEIILAE 227
RESULT 2
AD100558
ID AD100558 standard; protein; 328 AA.
XX
AC AD100558;
XX
DT 22-APR-2004 (first entry)
XX
DE Human TANGO 197 HisTag fusion protein - plasmid pO615.
XX
KW fusion; von Willebrand factor A-like domain; vWF; antibacterial;
KW cutaneous; inhalation anthrax; human; TANGO 197 HisTag fusion; mutant;
KW plasmid pO615; mutcin.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN US2003144193-A1.
XX
PD 31-JUL-2003.
XX
PF 24-JUL-2002; 2002US-00201292.
XX
PR 20-DEC-2001; 2001US-00038307.
XX
PA (ROTT/) ROTTMAN J B.
PA (OKEE/) O'KEEFE T L.
PA (OZKA/) OZKAYMAK E.
PA (HEAL/) HEALEY J J.
XX
PI Rottman JB, O'keefe TL, Ozkaymak E, Healey JJ;
XX WPI; 2003-720708/68.
XX
```

```
DR N-PSDB; AD100557.
XX
XX New TANGO 197 and/or TANGO 216 fusion polypeptides for preventing or
PT ameliorating symptoms of anthrax comprises a von Willebrand factor A-like
PT domain (vWF) amino acid sequence and an amino acid sequence heterologous
PT to the vWF.
XX
XX Claim 45; SEQ ID NO 26; 86pp; English.
XX
XX The invention relates to a novel fusion polypeptide comprising a von
CC Willebrand factor A-like domain (vWF) amino acid sequence and an amino
CC acid sequence heterologous to the vWF. The polypeptide of the invention
CC demonstrates antibacterial activities whilst the composition and method
CC may be useful in preventing or ameliorating the symptoms of cutaneous
CC and/or inhalation anthrax. The current sequence is that of the human
CC TANGO 197 HisTag fusion protein of the invention.
XX
SQ Sequence 328 AA;
Query Match 100.0%; Score 970; DB 7; Length 328;
Best Local Similarity 100.0%; Pred. No. 1.8e-99;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGFPLVFLDKSGSVLHHNNEIYFVPEQLAHKFTSPQLRMSFIYFSTRGTTLMKLTEDRE 60
DB 41 GGFPLVFLDKSGSVLHHNNEIYFVPEQLAHKFTSPQLRMSFIYFSTRGTTLMKLTEDRE 100
QY 61 QIRGSELEOKVLPGGDTYHMEGFERRASEQIYYENRQGYRTASVITALTDEGLHEDLFFY 120
DB 101 QIRGSELEOKVLPGGDTYHMEGFERRASEQIYYENRQGYRTASVITALTDEGLHEDLFFY 160
QY 121 SEREANSRDLGAIYVCVGVDFNETQLARIADSKDHVFVNDGFQALGGIHSILKKSC 180
DB 161 SEREANSRDLGAIYVCVGVDFNETQLARIADSKDHVFVNDGFQALGGIHSILKKSC 220
QY 181 IEIILAE 187
DB 221 IEIILAE 227
RESULT 3
ADM64584
ID ADM64584 standard; protein; 328 AA.
XX
AC ADM64584;
XX
DT 03-JUN-2004 (first entry)
XX
DE Human TANGO197-His tag fusion protein #2.
XX
KW antibacterial; gene therapy;
KW von Willebrand factor A-like domain amino acid sequence;
KW vWF amino acid sequence; anthrax; Bacillus anthracis; cutaneous anthrax;
KW inhalation anthrax; human; TANGO197; his tag; fusion protein.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN US2003134786-A1.
XX
PD 17-JUL-2003.
XX
PF 20-DEC-2001; 2001US-00038307.
XX
PR 20-DEC-2001; 2001US-00038307.
XX
PA (ROTT/) ROTTMAN J B.
PA (OKEE/) O'KEEFE T L.
PA (OZKA/) OZKAYMAK E.
PA (HEAL/) HEALEY J J.
XX
PI Rottman JB, O'keefe TL, Ozkaymak E, Healey JJ;
XX
```

DR WPI, 2003-829643/77.
DR N-PSDB; ADM64583.
PT New TANGO 197 and/or TANGO 216 fusion polypeptides for preventing or
PT ameliorating symptoms of anthrax comprises a von Willebrand factor A-like
PT domain (WVF) amino acid sequence and an amino acid sequence heterologous
PT to the WVF.
XX
XX
PS Claim 44; SEQ ID NO 26; 64pp; English.
CC The invention describes a fusion polypeptide comprising a von Willebrand
CC factor A-like domain (WVF) amino acid sequence and an amino acid sequence
CC heterologous to the WVF. Also described are: a method of preventing or
CC ameliorating a symptom of anthrax in a subject thought to be at risk for
CC exposure to or suspected of having been exposed to *Bacillus anthracis*;
CC and a pharmaceutical composition comprising the novel fusion polypeptide.
CC The composition and method are useful in preventing or ameliorating
CC symptoms of cutaneous and/or inhalation anthrax. This is the amino acid
CC sequence of a fusion protein comprising mature human TANGO197, thrombin
CC cleavage site and his tag that can be used to treat exposure to or
CC prevent a symptom of anthrax.
XX
SQ Sequence 328 AA;
Query Match 100.0%; Score 970; DB 7; Length 328;
Best Local Similarity 100.0%; Pred. No. 1.8e-99;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGFPLVFLDKSGSVLHNMNIYYFVEQLAHKFTSPQLRMSFIVFSTRGTTMLKLTEDRE 60
DB 41 GGFPLVFLDKSGSVLHNMNIYYFVEQLAHKFTSPQLRMSFIVFSTRGTTMLKLTEDRE 100
QY 61 QIRGLEBELQKVLPGSGTYMHGFERASEQIYYENRQGYRTASVITALTGELHEDLFFY 120
DB 101 QIRGLEBELQKVLPGSGTYMHGFERASEQIYYENRQGYRTASVITALTGELHEDLFFY 160
QY 121 SERANRSRDLGAIIVCVGVKDFNETQARIADSKDHFVPVNDGFQALOGIHSILKSC 180
DB 161 SERANRSRDLGAIIVCVGVKDFNETQARIADSKDHFVPVNDGFQALOGIHSILKSC 220
QY 181 IEILAAE 187
DB 221 IEILAAE 227
RESULT 4
AAB01422 standard; protein; 333 AA.
XX ID AAB01422 standard; protein; 333 AA.
XX AC AAB01422;
XX DT 20-OCT-2000 (first entry)
XX DE Human TANGO 197.
XX KM TANGO; 128; 140; 197; 212; 213; 224; 239; modulating agent; asthma;
XX graft versus-host diseases; rheumatoid arthritis; psoriasis;
XX inflammatory bowel disease; septic shock; ulcerative colitis;
XX Crohn's disease; chronic myelogenous leukemia; cancer; liver disease;
XX Hodgkin's disease; osteoarthritis; Lyme's disease; cachexia;
XX autoimmune disease; myasthenia gravis; autoimmune diabetes;
XX systemic lupus erythematosus; transgenic animal; diagnosis; prognosis;
XX prophylactic; therapeutic; human.
XX OS Homo sapiens.
XX PN MO200039284-A1.
XX PD 06-JUL-2000.
XX PF 23-DEC-1999; 99MO-US031025.
XX PR 30-DEC-1998; 98US-00223546.

XX (MILL-) MILLENNIUM PHARM INC.
PA Holtzman DA;
XX WPI, 2000-465743/40.
XX DR N-PSDB; AAA47455.
XX
XX
PT Novel nucleic acid sequences encoding TANGO-128, 140, 197, 212, 213, 224
PT and 239 polypeptides useful for the treatment of asthma, rheumatoid
PT arthritis, psoriasis and autoimmune diseases.
XX
XX
PS Claim 8; Fig 4; 209pp; English.
CC Nucleic acids encoding TANGO polypeptides are useful as modulating agents
CC for regulating cellular processes like asthma, graft versus-host
CC diseases, rheumatoid arthritis, psoriasis, inflammatory bowel disease,
CC septic shock, ulcerative colitis, Crohn's disease, chronic myelogenous
CC leukemia, cancer, liver disease, Hodgkin's disease, osteoarthritis,
CC Lyme's disease, cachexia and autoimmune diseases e.g. myasthenia gravis,
CC autoimmune diabetes and systemic lupus erythematosus. The nucleic acids
CC are also useful for producing transgenic animals and the TANGO
CC polypeptides themselves. Partial TANGO-128, 140, 197, 212, 213, 224, 239
CC sequences are useful in forensic biology, for diagnostic assays,
CC prognostic assays, pharmacogenomics and for monitoring clinical trials.
CC TANGO polypeptides are suitable for both prophylactic and therapeutic
CC methods for treating a subject at risk of a disorder or having a disorder
CC associated with aberrant TANGO expression. A wide range of cellular
CC disorders can be treated
XX
SQ Sequence 333 AA;
Query Match 100.0%; Score 970; DB 3; Length 333;
Best Local Similarity 100.0%; Pred. No. 1.9e-99;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGFPLVFLDKSGSVLHNMNIYYFVEQLAHKFTSPQLRMSFIVFSTRGTTMLKLTEDRE 60
DB 41 GGFPLVFLDKSGSVLHNMNIYYFVEQLAHKFTSPQLRMSFIVFSTRGTTMLKLTEDRE 100
QY 61 QIRGLEBELQKVLPGSGTYMHGFERASEQIYYENRQGYRTASVITALTGELHEDLFFY 120
DB 101 QIRGLEBELQKVLPGSGTYMHGFERASEQIYYENRQGYRTASVITALTGELHEDLFFY 160
QY 121 SERANRSRDLGAIIVCVGVKDFNETQARIADSKDHFVPVNDGFQALOGIHSILKSC 180
DB 161 SERANRSRDLGAIIVCVGVKDFNETQARIADSKDHFVPVNDGFQALOGIHSILKSC 220
QY 181 IEILAAE 187
DB 221 IEILAAE 227
RESULT 5
ABP54905 standard; protein; 333 AA.
XX ID ABP54905 standard; protein; 333 AA.
XX AC ABP54905;
XX DT 08-JAN-2003 (first entry)
XX DE Human anthrax toxin receptor.
XX KM Anthrax; toxin; receptor; human; antibacterial.
XX OS Homo sapiens.
XX PN
XX Key Location/Qualifiers
XX FT 1..27
XX FT Peptide /label= Signal_peptide
XX FT Protein 28..333
XX FT /label= Mature_protein

PN WO200246228-A2.
XX
XX
PD 13-JUN-2002.
XX
XX
PF 03-OCT-2001, 2001WO-US030941.
XX
XX
PR 05-DEC-2000, 2000US-0251481P.
XX
XX (WISC) WISCONSIN ALUMNI RES FOUND.
XX
XX
PI Young JAT, Bradley KA, COLLIER RJ, Mogridge JS;
XX
XX
XX WPI; 2002-713235/77.
DR
DR N-PSDB; ABV73863.
XX
XX
XX Novel isolated polypeptide useful for identifying agent that prevents or
PT reduces effect of anthrax toxin on host cell, for treating human or non-
PT human animal suffering from anthrax.
XX
XX
XX Claim 1, Page 42-43; 45pp; English.

CC The present sequence is the protein sequence of a polypeptide identified
 CC as a human anthrax toxin receptor (ATR) polypeptide on the basis of
 CC identity to a newly isolated human ATR (see AB54903). The 2 polypeptides
 CC are identical between amino acids 1-317, but differ thereafter at the C-
 CC terminus. The present polypeptide, encoded by IMAGE CLONE 4563020, has no
 CC previously known function, and there has been no prior indication that it
 CC is a complete or partial ATR. The invention provides ATR polypeptides and
 CC polynucleotides, vectors, host cells, and transgenic and knock-out
 CC animals. It also provides methods for identifying molecules that bind the
 CC ATR and which reduce the toxicity of anthrax toxin. A claimed method for
 CC treating anthrax in a human or animal involves administering an agent
 CC that inhibits binding between anthrax toxin protective antigen (PA) and
 CC ATR at a level effective to reduce the severity of anthrax. Suitable
 CC agents include the present polypeptide or a PA-binding fragment of it, a
 CC PA-binding polypeptide at least 80% identical to these, a fusion protein,
 CC a monoclonal or polyclonal antibody, a polysaccharide, a lipid or a
 CC nucleic acid
 CC
 CC
 CC Sequence 333 AA:

Query Match	100.0%;	Score 970;	DB 5;	Length 333;
Best Local Similarity	100.0%;	Pred. No. 1.9e-99;		
Matches 187;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	GGFPLYIILKSSVLHNNHNIYYFEOLAHKFIISPOLMSTIVSTGTLMLKTEDE	60
Db	41 GGFPLYIILKSSVLHNNHNIYYFEOLAHKFIISPOLMSTIVSTGTLMLKTEDE	100
Qy	61 QIRQSGEELQKVLPGSDTYMHGFEFASQIYYENNQGYRTASVILALTDGELHEDLFFY	120
Db	101 QIRQSGEELQKVLPGSDTYMHGFEFASQIYYENNQGYRTASVILALTDGELHEDLFFY	160
Qy	121 SERAANRSRLGAIYVCVGYKDNFNEQLARIADSKDHVPPVNDGFGQALQGIIHSLKSC	180
Db	161 SERAANRSRLGAIYVCVGYKDNFNEQLARIADSKDHVPPVNDGFGQALQGIIHSLKSC	220
Qy	181 IEILAAE 187	
Db	221 IEILAAE 227	

RESULT 6
ADI00534
ID ADI00534 standard; protein; 333 AA.

AC ADI00534;
XX
DT 22-APR-2004 (first entry)
XX
DE Human TANGO 197 protein.
XX
FM fusion; von Willebrand factor A-like domain; vWF, antibacterial;

KW	cutaneous; inhalation anthrax;	human; TANGO 197
XX		
OS	Homo sapiens.	
XX		
PN	US2003144193-A1.	
XX		
PD	31-JUL-2003.	

PF	24-JUL-2002; 2002US-00201292.
XX	
PR	20-DEC-2001; 2001US-00038307.

PA (ROTT/) ROTTMAN J B.
PA (OKEE/) O'KEEFE T L.
PA (OZKA/) OZKAYNAK E.
PA (HEAL/) HEALEY J J.

PI Rottman JB, O'Keefe TL, Ozkaynak E, Healey JJ;

DR WPI; 2003-720708/68.
DR N-PSDB; ADI00533.

PT New TANGO 197 and/or TANGO 216 fusion polypeptides for preventing or ameliorating symptoms of anthrax comprises a von Willebrand factor A-like domain (vWF) amino acid sequence and an amino acid sequence heterologous to the vWF.

PS Claim 26; SEQ ID NO 2; 86pp; English.

CC The invention relates to a novel fusion polypeptide comprising a von
CC Willbrand factor A-like domain (VWF) amino acid sequence and an amino
CC acid sequence heterologous to the VWF. The polypeptide of the invention
CC demonstrates antibacterial activities whilst the composition and method
CC may be useful in preventing or ameliorating the symptoms of cutaneous
CC and/or intraocular antrax. The current sequence is that of the human
CC TANGO 197 protein of the invention.

SQ Sequence 333 AA;

Query Match	100.0%	Score 970 ; DB 7 ;	Length 333 ;
Best Local Similarity	100.0%	Pred. No. 1, 9e-99 ;	
Matches 187 ;	Conservative 0 ;	Mismatches 0 ;	Indels 0 ; Gaps 0 ;

QY	1	GGFDYFLINDSGSVLHHMNEIYYVEVQLAHFISPOLMSFVFTRTGTLAKLTEDRE	60
Db	41	GGFDYFLINDSGSVLHHMNEIYYVEVQLAHFISPOLMSFVFTRTGTLAKLTEDRE	1000
QY	61	QIRGLELQKVLPGGDTYYMHBGFERASQIYYENRQGYRTASVIALTDGELHEDLFFY	1200
Db	101	QIRGLELQKVLPGGDTYYMHBGFERASQIYYENRQGYRTASVIALTDGELHEDLFFY	1600
QY	121	SRENRNRSDGAIYYCVGVKDFNETQLARIADSKDHFVPVNDGFQALQGIHSLTKSC	1800
Db	161	SRENRNRSDGAIYYCVGVKDFNETQLARIADSKDHFVPVNDGFQALQGIHSLTKSC	2200
QY			
		181 ITILAAE 187	
Db	221	ITILAAE 227	

RESULT 7	
ADM64568	
ID	ADM64568 standard; protein; 333 AA

AC ADM64568;

DT 03-JUN-2004 (first entry)

Human von Willebrand factor A-like domain protein TANGO197.

KW antibacterial; gene therapy;
KW von Willebrand factor A-like domain amino acid sequence;
KW VWF amino acid sequence; anthrax; *Bacillus anthracis*; cutaneous anthrax;

KM		Inhalation anthrax; human; TANGO197.
XX		
OS	Homo sapiens.	
XX		
PN	US2003134786-A1.	
XX		
PD	17-JUL-2003.	
XX		
PF	20-DEC-2001; 2001US-00038307.	
XX		
PR	20-DEC-2001; 2001US-00038307.	
XX		
PA	(ROTT/) ROTTMAN J B.	
PA	(OKEE/) O'KEEFE T L.	
PA	(OZKA/) OZKAYNAK E.	
PA	(HEAL/) HEALEY J J.	
PI	Rottman JB, O'keefe TL, Ozkaynak E, Healey JJ;	
DR	WPI; 2003-829643/77.	
N-FSDB; ADW64567.		
PT	New TANGO 197 and/or TANGO 216 fusion polypeptides for preventing or ameliorating symptoms of anthrax comprises a von Willebrand factor A-like domain (vWF) amino acid sequence and an amino acid sequence heterologous to the vwf.	
PS	Claim 26; SEQ ID NO 2; 64pp; English.	
CC	The invention describes a fusion polypeptide comprising a von Willebrand factor A-like domain (vWF) amino acid sequence and an amino acid sequence heterologous to the vwf. Also described are: a method of preventing or ameliorating a symptom of anthrax in a subject thought to be at risk for exposure to or suspected of having been exposed to Bacillus anthracis; and a pharmaceutical composition comprising the novel fusion polypeptide. The composition and method are useful in preventing or ameliorating symptoms of cutaneous and/or inhalation anthrax. This is the amino acid sequence of a human von Willebrand factor A-like domain (vWF) amino acid sequence TANG0197.	
SQ	Sequence 333 AA;	
Query Match	100.0%; Score 970; DB 7; Length 333;	
Best Local Similarity	100.0%; Pred. No. 1.9e+99;	
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0		
Dd	1 GGFDLYFLIDSGSVLHMHNEIYYFVEQLAHKFISPOLMSFIIVSTRTGTLMKLTEDRE 60 41 GGFDLYFLIDSGSVLHMHNEIYYFVEQLAHKFISPOLMSFIIVSTRTGTLMKLTEDRE 100	
Gy	61 QIRQLESLQVLPGGDVTVMHGFFRRASRQIYYENRGRTASVIALTDSGLIHDLDFPY 120 101 QIRQLESLQVLPGGDVTVMHGFFRRASRQIYYENRGRTASVIALTDSGLIHDLDFPY 160	
Dd	121 SEREANRDICGAIIYCVGVDPFNETOLARIADSKDHYPVNVDGFQALOGIHTSIKKSC 180 161 SEREANRDICGAIIYCVGVDPFNETOLARIADSKDHYPVNVDGFQALOGIHTSIKKSC 220	
Gy	181 TEILAAE 187 	
Dd	221 TEILAAE 227	
RESULT 8		
ID	ADI00554	
AC	ADI00554 standard; protein; 342 AA.	
DT	22-APR-2004 (first entry)	
XN	Human TANGO 197 FLAG fusion protein - plasmid p0613.	

Query Match	Best Local Similarity	100.0%	Score 970	DB 7	Length 342
Matches 187	Conservative 0	Mismatches 0	Indels 0	Gaps 0	
QY 1	GGFDLYPLIDKSGSVLHNMNIYFFVEQLAKHFI	SPQLRMSFI	FSTRGTLMLKLTEDRE	60	
DB 50	GGFDLYPLIDKSGSVLHNMNIYFFVEQLAKHFI	SPQLRMSFI	FSTRGTLMLKLTEDRE	109	
QY 61	QIRQGLELQVLPQGGDTVMHEGFERAS	EQIYYENRQ	GRTASVIALTDGELHEDLFFY	120	
DB 110	QIRQGLELQVLPQGGDTVMHEGFERAS	EQIYYENRQ	GRTASVIALTDGELHEDLFFY	169	
QY 121	SERFANSRDLGAIYVCVGVDPFNETOLARI	ADSKDHVFP	NDGFOALOGIHSILKKS	180	
DB 170	SERFANSRDLGAIYVCVGVDPFNETOLARI	ADSKDHVFP	NDGFOALOGIHSILKKS	229	
QY 181	IEIILAE 187				
DB 230	IEIILAE 236				
RESULT 9					
ID ADM64580					
XX ADM64580	standard; protein; 342 AA.				
AC ADM64580;					
XX 03-JUN-2004	(first entry)				
XX DE	Human TANGO197-FLAG epitope fusion protein.				

KM antibacterial; gene therapy;
 KM von Willebrand factor A-like domain amino acid sequence;
 KM vWF amino acid sequence; anthrax; Bacillus anthracis; cutaneous anthrax;
 KM inhalation anthrax; human; TANGO197; FLAG; fusion protein; gene.
 XX Homo sapiens.
 OS Synthetic.
 XX US2003134786-A1.
 XX 17-JUL-2003.
 XX 20-DEC-2001; 2001US-00038307.
 XX 20-DEC-2001; 2001US-00038307.
 XX 20-DEC-2001; 2001US-00038307.
 XX (ROT/) ROTTMAN J B.
 PA (OKEE/) O'KEEFE T L.
 PA (OZKA/) OZKAYNAK E.
 PA (HEAL/) HEALEY J J.
 XX Rotman JB, O'Keefe TL, Ozkaynak E, Healey JJ;
 PI WPI; 2003-829643/77.
 DR N-PSDB; ADM64579.
 XX New TANGO 197 and/or TANGO 216 fusion polypeptides for preventing or
 PT ameliorating symptoms of anthrax comprises a von Willebrand factor A-like
 PT domain (vWF) amino acid sequence and an amino acid sequence heterologous
 PT to the vWF.
 XX
 XX Claim 44; SEQ ID NO 22; 64pp; English.
 XX
 CC The invention describes a fusion polypeptide comprising a von Willebrand
 CC factor A-like domain (vWF) amino acid sequence and an amino acid sequence
 CC heterologous to the vWF. Also described are: a method of preventing or
 CC ameliorating a symptom of anthrax in a subject thought to be at risk for
 CC exposure to or suspected of having been exposed to Bacillus anthracis;
 CC and a pharmaceutical composition comprising the novel fusion polypeptide.
 CC The composition and method are useful in preventing or ameliorating
 CC symptoms of cutaneous and/or inhalation anthrax. This is the amino acid
 CC sequence of a fusion protein comprising human TANGO197 and FLAG that can
 CC be used to treat exposure to or prevent a symptom of anthrax.
 CC
 XX Sequence 342 AA;
 SQ
 Query Match 100.0%; Score 970; DB 7; Length 342;
 Best Local Similarity 100.0%; Pred. No. 1.9e-99;
 Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGFDFLFDKSGSVLHNMNIYFVEQLAHKFIISPOLRMSFIVFSTRGTTLMKLTEDRE 60
 Db 50 GGFDFLFDKSGSVLHNMNIYFVEQLAHKFIISPOLRMSFIVFSTRGTTLMKLTEDRE 109
 QY 61 QIRGLEELQKVLPGGDTYHBEGERASEQIYYENRQGYRTASVIALTDGELHEDLFFY 120
 Db 110 QIRGLEELQKVLPGGDTYHBEGERASEQIYYENRQGYRTASVIALTDGELHEDLFFY 169
 QY 121 SEREANKSRDLGAIYVCVGVDFNETOLARIADSKDHVFPVNDGFQALQGIHISILKSC 180
 Db 170 SEREANKSRDLGAIYVCVGVDFNETOLARIADSKDHVFPVNDGFQALQGIHISILKSC 229
 QY 181 IEILAAE 187
 Db 230 IEILAAE 236
 RESULT 10
 ADI00556
 ID ADI00556 standard; protein; 345 AA.
 XX AC ADI00556;
 XX

DT 22-APR-2004 (first entry)
 XX
 DE Human TANGO 197 Histag fusion protein - plasmid p0614.
 XX
 KM fusion; von Willebrand factor A-like domain; vWF; antibacterial;
 KM cutaneous; inhalation anthrax; human; TANGO 197 Histag fusion; mutant;
 KM plasmid p0614; mutcin.
 XX
 XX Homo sapiens.
 OS Synthetic.
 XX US2003144193-A1.
 XX 31-JUL-2003.
 XX 24-JUL-2002; 2002US-00201292.
 XX 20-DEC-2001; 2001US-00038307.
 XX (ROT/) ROTTMAN J B.
 PA (OKEE/) O'KEEFE T L.
 PA (OZKA/) OZKAYNAK E.
 PA (HEAL/) HEALEY J J.
 XX Rotman JB, O'Keefe TL, Ozkaynak E, Healey JJ;
 PI WPI; 2003-720708/68.
 DR N-PSDB; ADI00555.
 XX New TANGO 197 and/or TANGO 216 fusion polypeptides for preventing or
 PT ameliorating symptoms of anthrax comprises a von Willebrand factor A-like
 PT domain (vWF) amino acid sequence and an amino acid sequence heterologous
 PT to the vWF.
 XX
 XX Claim 45; SEQ ID NO 24; 86pp; English.
 XX
 CC The invention relates to a novel fusion polypeptide comprising a von
 CC Willebrand factor A-like domain (vWF) amino acid sequence and an amino
 CC acid sequence heterologous to the vWF. The polypeptide of the invention
 CC demonstrates antibacterial activities whilst the composition and method
 CC may be useful in preventing or ameliorating the symptoms of cutaneous
 CC and/or inhalation anthrax. The current sequence is that of the human
 CC TANGO 197 Histag fusion protein of the invention.
 CC
 XX Sequence 345 AA;
 SQ
 Query Match 100.0%; Score 970; DB 7; Length 345;
 Best Local Similarity 100.0%; Pred. No. 2e-99;
 Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGFDFLFDKSGSVLHNMNIYFVEQLAHKFIISPOLRMSFIVFSTRGTTLMKLTEDRE 60
 Db 41 GGFDFLFDKSGSVLHNMNIYFVEQLAHKFIISPOLRMSFIVFSTRGTTLMKLTEDRE 100
 QY 61 QIRGLEELQKVLPGGDTYHBEGERASEQIYYENRQGYRTASVIALTDGELHEDLFFY 120
 Db 101 QIRGLEELQKVLPGGDTYHBEGERASEQIYYENRQGYRTASVIALTDGELHEDLFFY 160
 QY 121 SEREANKSRDLGAIYVCVGVDFNETOLARIADSKDHVFPVNDGFQALQGIHISILKSC 180
 Db 161 SEREANKSRDLGAIYVCVGVDFNETOLARIADSKDHVFPVNDGFQALQGIHISILKSC 220
 QY 181 IEILAAE 187
 Db 221 IEILAAE 227
 RESULT 11

ID	ADM64582
ID	ADM64582 standard; protein; 345 AA.
XX	
AC	ADM64582;
XX	
DT	03-JUN-2004 (first entry)
XX	
DE	Human TANGO197-His tag fusion protein #1.
XX	
KV	antibacterial; gene therapy;
KW	von Willebrand factor A-like domain amino acid sequence;
KM	von Willebrand factor A-like domain; anthrax; Bacillus anthracis; cutaneous anthrax;
XX	inhalation anthrax; human; TANGO197; his tag; fusion protein; gene.
OS	Homo sapiens.
XX	Synthetic.
PX	US2003134786-A1.
PD	17-JUL-2003.
PF	20-DEC-2001; 2001US-00038307.
PR	20-DEC-2001; 2001US-00038307.
PA	(ROTT/) ROTTMAN J B.
PA	(OKEE/) O'KEEFE T L.
PA	(OZKA/) OZKAYNAK E.
PA	(HEAL/) HEALEY J J.
FL	Rottman JB, O'Keefe TL, Ozkaynak E, Healey JJ;
DR	WPI; 2003-829643/77.
DR	N-Psdb; ADM64581.
PT	New TANGO 197 and/or TANGO 216 fusion polypeptides for preventing or
PT	ameliorating symptoms of anthrax comprises a von Willebrand factor A-like
PT	domain (vWF) amino acid sequence and an amino acid sequence heterologous
PT	to the vWF.
PS	Claim 44; SEQ ID NO 24; 64pp; English.
XX	
CC	The invention describes a fusion polypeptide comprising a von Willebrand
CC	factor A-like domain (vWF) amino acid sequence and an amino acid sequence
CC	heterologous to the vWF. Also described are: a method of preventing or
CC	ameliorating a symptom of anthrax in a subject thought to be at risk for
CC	exposure to or suspected of having been exposed to Bacillus anthracis;
CC	and a pharmaceutical composition comprising the novel fusion polypeptide.
CC	The composition and method are useful in preventing or ameliorating
CC	symptoms of cutaneous and/or inhalation anthrax. This is the amino acid
CC	sequence of a fusion protein comprising human TANGO197, thrombin cleavage
CC	site and his tag that can be used to treat exposure to or prevent a
CC	symptom of anthrax.
XX	
SQ	Sequence 345 AA;
Query Match	100.0%; Score 970; DB 7; Length 345;
Best Local Similarity	100.0%; Pred. No. 2e-99; Indels 0; Gaps 0
Matches 187; Conservative	0; Mismatches 0;
Dy	1 GGFDLYFLIDSGSVLAHMHNEIYYVEQLAHKFISPOLMSFIVSTRGTIMKLTEORE 60
Dd	41 GGFPLYFLIDSGSVLAHMHNEIYYVEQLAHKFISPOLMSFIVSTRGTIMKLTEORE 100
Dy	61 QIRQGLBELQKYLPGGDIYMEHGFRASFOIYYENRQGRTASVIIALTDSGLHEDLFY 120
Dd	101 QIRQGLBELQKYLPGGDIYMEHGFRASFOIYYENRQGRTASVIIALTDSGLHEDLFY 160
Dy	121 SERANRRBDIGAIYYCVGVDFNFTOLARINDSDHPVPVNDGFOALOGIIHSILTKSC 180
Dd	161 SERANRRBDIGAIYYCVGVDFNFTOLARINDSDHPVPVNDGFOALOGIIHSILTKSC 220
Dy	181 IETILAEE 187

Db	221	IEILAAE	227
RESULT 12	ABP54903		
ID	ABP54903	standard; protein; 368 AA.	
XX	AC	ABP54903;	
XX	DT	08-JAN-2003 (first entry)	
XX	DE	Human anthrax toxin receptor.	
XX	XX	Anthrax; toxin; receptor; human; antibacterial.	
XX	OS	Homo sapiens.	
XX	Key	Location/qualifiers	
FT	Peptide	1..27	
FT		/label= signal_peptide	
FT	Region	27..321	
FT		/note= "PR-binding fragment, begins at any amino acid in the range 27-43 and ends at any amino acid in the range 221-321, region specifically described in Claim 3"	
FT	Protein	28..368	
FT		/label= Mature_protein	
FT	Domain	28..320	
FT		/note= "extracellular domain"	
FT	Domain	44..216	
FT		/note= "von Willebrand factor A domain"	
FT	Region	50	
FT		/note= "forms metal ion-dependent adhesion site (MIDAS) motif with amino acid residues 52, 54, 118 and 150"	
FT	Region	52	
FT		/note= "forms metal ion-dependent adhesion site (MIDAS) motif with amino acid residues 50, 54, 118 and 150"	
FT	Region	54	
FT		/note= "forms metal ion-dependent adhesion site (MIDAS) motif with amino acid residues 50, 52, 118 and 150"	
FT	Region	118	
FT		/note= "forms metal ion-dependent adhesion site (MIDAS) motif with amino acid residues 50, 52, 54 and 150"	
FT	Region	150	
FT		/note= "forms metal ion-dependent adhesion site (MIDAS) motif with amino acid residues 50, 52, 54 and 150"	
FT	Domain	320..343	
FT		/note= "putative transmembrane domain"	
FT	Domain	344..368	
FT		/note= "cytoplasmic domain"	
XX	MO200246228-A2.		
XX	13-JUN-2002.		
XX	03-OCT-2001; 2001MO-US030941.		
XX	05-DEC-2000; 2000US-0251481P.		
XX	(WISC) WISCONSIN ALUMNI RES FOUND.		
XX	Young JAT, Bradley KA, Collier RJ, Mogridge JS;		
XX	WPI; 2002-713235/77.		
XX	N-PSDB; ABV73881.		
XX	Novel isolated polypeptide useful for identifying agent that prevents or reduces effect of anthrax toxin on host cell, for treating human or non-human animal suffering from anthrax.		
XX	Claim 1; Page 29-30; 45pp; English.		
XX	The present sequence is the protein sequence of a human surface-bound		

CC anthrax toxin receptor (ATR), as predicted from an isolated cDNA clone.
 CC Anthrax toxin protective antigen (PA) binds to the ATR at a von
 CC Willebrand factor A domain located in the extracellular domain of ATR.
 CC The invention provides ATR polypeptides and polynucleotides, vectors,
 CC host cells, and transgenic and knock-out animals. It also provides the
 CC methods for identifying molecules that bind the ATR and which reduce the
 CC toxicity of anthrax toxin. A claimed method for treating anthrax in a
 CC human or animal involves administering an agent that inhibits binding
 CC between PA and ATR at a level effective to reduce the severity of
 CC anthrax. Suitable agents include ATR or a PA-binding fragment of ATR, a
 CC PA-binding polypeptide at least 80% identical to these, a fusion protein,
 CC a monoclonal or polyclonal antibody, a polysaccharide, a lipid or a
 CC nucleic acid
 CC
 XX Sequence 368 AA;
 SQ
 Query Match 100.0%; Score 970; DB 5; Length 368;
 Best Local Similarity 100.0%; Pred. No. 2.1e-99;
 Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGFPLVFIIDKSGSVLHHNNEIYFVEQLAHKFIISPOLRMSFVFSRTGTLTKLTEDRE 60
 DB 41 GGFPLVFIIDKSGSVLHHNNEIYFVEQLAHKFIISPOLRMSFVFSRTGTLTKLTEDRE 100
 QY 61 QIRGLEELQKVLPGGDTYHMEGFERSAQIYENRQGYRTASVITALTGELHEDLFFY 120
 DB 101 QIRGLEELQKVLPGGDTYHMEGFERSAQIYENRQGYRTASVITALTGELHEDLFFY 160
 QY 121 SERANRSRDIGAIYVCVGVDPFNETOLARIADSKDHFVPVNDGFOLQGIHSLKSC 180
 DB 161 SERANRSRDIGAIYVCVGVDPFNETOLARIADSKDHFVPVNDGFOLQGIHSLKSC 220
 QY 181 IEILAAE 187
 DB 221 IEILAAE 227
 RESULT 13
 ADM64586
 ID ADM64586 standard; protein; 384 AA.
 AC ADM64586;
 XX
 XX 03-JUN-2004 (first entry)
 DE TANGO197 extracellular domain-mutant IgG Fc fusion protein #1.
 XX
 KW antibacterial; gene therapy;
 KW von Willebrand factor A-like domain amino acid sequence;
 KW vWF amino acid sequence; anthrax; Bacillus anthracis; cutaneous anthrax;
 KW inhalation anthrax; human; TANGO197; immunoglobulin G; IgG;
 KW fragment of crystallisation; Fc.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 XX US2003134786-A1.
 PN
 XX 17-JUL-2003.
 PD
 XX 20-DEC-2001; 2001US-00038307.
 PF
 XX 20-DEC-2001; 2001US-00038307.
 PR
 XX
 XX (ROTT/) ROTTMAN J B.
 PA (OKEE/) O'KEEFE T L.
 PA (OZKA/) OZKAYNAK E.
 PA (HEAL/) HEALEY J J.
 XX
 PI Rottman JB, O'keefe TL, Ozkaynak E, Healey JJ;
 XX WPI; 2003-829643/77.
 DR N-PSDB; ADM64585.
 DR

XX New TANGO 197 and/or TANGO 216 fusion polypeptides for preventing or
 PT ameliorating symptoms of anthrax comprises a von Willebrand factor A-like
 PT domain (vWF) amino acid sequence and an amino acid sequence heterologous
 PT to the vWF.
 XX
 XX Claim 44; SEQ ID NO 10; 64pp; English.
 XX
 CC The invention describes a fusion polypeptide comprising a von Willebrand
 CC factor A-like domain (vWF) amino acid sequence and an amino acid sequence
 CC heterologous to the vWF. Also described are: a method of preventing or
 CC ameliorating a symptom of anthrax in a subject thought to be at risk for
 CC exposure to or suspected of having been exposed to Bacillus anthracis;
 CC and a pharmaceutical composition comprising the novel fusion polypeptide.
 CC The composition and method are useful in preventing or ameliorating
 CC symptoms of cutaneous and/or inhalation anthrax. This is the amino acid
 CC sequence of a fusion protein comprising human TANGO197 extracellular
 CC region, minus the DG residues closest to the transmembrane region, and
 CC immunoglobulin G (196) fragment of crystallisation (Fc) with mutations
 CC L235A and G237A.
 CC
 XX Sequence 384 AA;
 SQ
 Query Match 100.0%; Score 970; DB 7; Length 384;
 Best Local Similarity 100.0%; Pred. No. 2.3e-99;
 Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGFPLVFIIDKSGSVLHHNNEIYFVEQLAHKFIISPOLRMSFVFSRTGTLTKLTEDRE 60
 DB 41 GGFPLVFIIDKSGSVLHHNNEIYFVEQLAHKFIISPOLRMSFVFSRTGTLTKLTEDRE 100
 QY 61 QIRGLEELQKVLPGGDTYHMEGFERSAQIYENRQGYRTASVITALTGELHEDLFFY 120
 DB 101 QIRGLEELQKVLPGGDTYHMEGFERSAQIYENRQGYRTASVITALTGELHEDLFFY 160
 QY 121 SERANRSRDIGAIYVCVGVDPFNETOLARIADSKDHFVPVNDGFOLQGIHSLKSC 180
 DB 161 SERANRSRDIGAIYVCVGVDPFNETOLARIADSKDHFVPVNDGFOLQGIHSLKSC 220
 QY 181 IEILAAE 187
 DB 221 IEILAAE 227
 RESULT 14
 AAE01439
 ID AAE01439 standard; protein; 403 AA.
 XX
 AC AAE01439;
 XX
 XX 17-JUL-2001 (first entry)
 DE Human gene 4 encoded secreted protein HMLPR02, SEQ ID NO:94.
 XX
 KW Human; secreted protein; proliferative disorder; cancer; tumour;
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
 KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
 KW cardiovascular disorder; angiogenic disorder; kidney disorder;
 KW gastrointestinal disorder; pregnancy-related disorder;
 KW endocrine disorder; infection; wound healing; vulnerability; cell culture;
 KW chemotaxis; food additive; gene therapy; binding partner identification;
 KW chromosome 19.
 XX
 OS Homo sapiens.
 OS
 XX
 XX Key 1.27
 FH Peptide /label= signal_peptide
 FT Protein 28.403
 FT /note= "Mature human secreted protein"

XX WO200134626-A1.
XX 17-MAY-2001.
XX 01-NOV-2000; 2000WO-US030045.
XX 05-NOV-1999; 99US-0163581P.
XX 30-JUN-2000; 2000US-0215133P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Komatsu GA, Moore PA, Biree CE, Ni J,
XX WPI; 2001-308778/32.
XX N-PSDB; AAD05303.
XX
XX New nucleic acid molecules encoding 28 human secreted proteins for
XX diagnosing, preventing, treating or ameliorating medical conditions and
XX used as food additives or preservatives.
XX
XX Claim 11; Page 485-486; 562pp; English.
XX
XX AAD05300-AAD05379 represent cDNAs corresponding to 28 human secreted
XX protein genes, and AAB01436-AAB01513 represent the proteins they encode.
XX AAB01514-AAB01544 represent human secreted protein fragments or variants.
XX The genes and their secreted proteins are useful for preventing, treating
XX or ameliorating medical conditions, e.g., by protein or gene therapy.
XX Pathological conditions can be diagnosed by determining the amount of the
XX new protein in a sample or by determining the presence of mutations in
XX the new genes. Specific uses are described for each of the 28 genes,
XX based on the tissues in which they are most highly expressed, and include
XX developing products for the diagnosis or treatment of proliferative
XX disorders, cancer, tumours, foetal and developmental abnormalities,
XX haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
XX diseases (e.g., rheumatoid arthritis), inflammation, allergies,
XX neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
XX cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
XX psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
XX angiotensin-related disorders, kidney disorders, gastrointestinal disorders,
XX pregnancy-related disorders, endocrine disorders, and infections. The
XX proteins can also be used to aid wound healing and epithelial cell
XX proliferation, to prevent skin aging due to sunburn, to maintain organs
XX before transplantation, for supporting cell culture of primary tissues,
XX to regenerate tissues, to identify their cognate ligands or binding
XX partners, and in chemotaxis, and can be used as a food additive or
XX preservative to modify storage properties. Antibodies specific for a
XX protein of the invention can be used in alleviating symptoms associated
XX with the disorders mentioned above, and in diagnostic immunoassays e.g.,
XX radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The
XX present sequence represents a human secreted protein of the invention
XX
XX Sequence 403 AA;
XX
XX Query Match 100.0%; Score 970; DB 4; Length 403;
XX Best Local Similarity 100.0%; Pred. No. 2,4e-99;
XX Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 GGFPLVFLDKSGSVLHNMNIYFVEQLAKHFIKSPOLRMSFIYFSTRGTTMLKLTEDRE 60
DB 41 GGFPLVFLDKSGSVLHNMNIYFVEQLAKHFIKSPOLRMSFIYFSTRGTTMLKLTEDRE 100
QY QIRGLELQKLVLCGGDTYMHGFERASEQIYYENRGQYRTASVIALTDGELHEDLPFY 120
DB 101 QIRGLELQKLVLCGGDTYMHGFERASEQIYYENRGQYRTASVIALTDGELHEDLPFY 160
QY 121 SERBANSRDIGAIYVCVGVDFMETQARIADSKDHVFPVNDGFOALQGIHSILKSC 180
DB 161 SERBANSRDIGAIYVCVGVDFMETQARIADSKDHVFPVNDGFOALQGIHSILKSC 220
QY 181 IETLAE 187
DB 221 IETLAE 227

RESULT 15
ABG63874
ID ABG63874 standard; protein, 403 AA.
XX
XX AC ABG63874;
XX
XX DT 27-AUG-2002 (first entry)
XX
XX DE Human albumin fusion protein #549.
XX
XX KM Albumin fusion protein, therapeutic protein X; human albumin; HA;
XX human serum albumin; HSA; cancer; reproductive disorder;
XX digestive disorder; immune disorder; endocrine disorder;
XX haematopoietic disorder; neural disorder; connective disorder;
XX cytosolic; antifertility; antiinflammatory; antitumor;
XX immunomodulator; anti-HIV; antidiabetic; hemostatic; neurotropic;
XX neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
XX osteopathic; antiarthritic.
XX
XX OS Homo sapiens.
XX OS Synthetic.
XX
XX WO200177137-A1.
XX
XX PD 18-OCT-2001.
XX
XX PE 12-APR-2001; 2001WO-US011988.
XX
XX PR 12-APR-2000; 2000US-0229358P.
XX PR 25-APR-2000; 2000US-0199384P.
XX PR 21-DEC-2000; 2000US-0256931P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX P1 Rosen CA, Haseltine WA;
XX
XX WPI; 2002-010886/01.
XX
XX PT New fusion protein for treating disease e.g. diabetes comprises an
XX albumin fused to a therapeutic protein.
XX
XX PS Claim 1; Page 874-875; 2102pp; English.
XX
XX CC The present invention relates to albumin fusion proteins comprising a
XX therapeutic protein X and human albumin (HA, also known as human serum
XX albumin, HSA). The proteins are useful for treating a disease or disorder
XX that may be modulated by therapeutic protein X. The albumin extends the
XX shelf-life of protein X, and may increase its biological in vitro/in vivo
XX activity. The protein is useful for treating and diagnosing disorders
XX such as cancer, reproductive disorders, digestive disorders (e.g. Crohn's
XX disease, ulcerative colitis), immune disorders (e.g. acquired
XX immunodeficiency syndrome, AIDS), endocrine disorders (e.g. diabetes),
XX haematopoietic disorders, neural disorders (e.g. Alzheimer's,
XX Parkinson's, Creutzfeldt-Jacob disease, encephalomyelitis, meningitis,
XX schizophrenia), and connective disorders (e.g. osteoporosis, arthritis).
XX ABG63326-ABG65518 represent albumin fusion proteins of the invention
XX
XX Sequence 403 AA;
XX
XX Query Match 100.0%; Score 970; DB 5; Length 403;
XX Best Local Similarity 100.0%; Pred. No. 2,4e-99;
XX Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 GGFPLVFLDKSGSVLHNMNIYFVEQLAKHFIKSPOLRMSFIYFSTRGTTMLKLTEDRE 60
DB 41 GGFPLVFLDKSGSVLHNMNIYFVEQLAKHFIKSPOLRMSFIYFSTRGTTMLKLTEDRE 100
QY QIRGLELQKLVLCGGDTYMHGFERASEQIYYENRGQYRTASVIALTDGELHEDLPFY 120
DB 101 QIRGLELQKLVLCGGDTYMHGFERASEQIYYENRGQYRTASVIALTDGELHEDLPFY 160

Qy 121 SEREANSRDI GAI VYCVGVKDNFNETQ LARIADSKDHVFPVNDGFQALOGI IHSILKSC 180
Db 161 SEREANSRDI GAI VYCVGVKDNFNETQ LARIADSKDHVFPVNDGFQALOGI IHSILKSC 220
Qy 181 IEIILAE 187
Db 221 IEIILAE 227

Search completed: June 13, 2005, 19:56:52
Job time : 71.2347 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: June 13, 2005, 19:46:02 ; Search time 18.054 Seconds
(without alignments)
773.200 Million cell updates/sec

Title: US-09-970-076-2_COPY_41_227

Perfect score: 1 GGFDFLYPLDKSGSVLHNM.....LQGIHSILKSCIEILAAE 187

Sequence: BLOSUM62

Scoring table: Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	142.5	14.7	435	5 PCT-US95-04439-1	Sequence 1, Appl
2	141.5	14.6	1151	1 US-08-286-889-37	Sequence 37, Appl
3	141.5	14.6	1151	1 US-08-485-618-37	Sequence 37, Appl
4	141.5	14.6	1151	1 US-08-362-652-37	Sequence 37, Appl
5	141.5	14.6	1151	2 US-08-605-672-37	Sequence 37, Appl
6	141.5	14.6	1151	2 US-08-482-293A-37	Sequence 37, Appl
7	141.5	14.6	1151	3 US-08-943-363-37	Sequence 37, Appl
8	141.5	14.6	1151	3 US-09-193-043-37	Sequence 37, Appl
9	141.5	14.6	1151	4 US-09-688-307A-37	Sequence 37, Appl
10	141.5	14.6	1151	4 US-09-350-259-37	Sequence 37, Appl
11	141.5	14.6	1161	1 US-08-485-618-55	Sequence 55, Appl
12	141.5	14.6	1161	1 US-08-362-652-55	Sequence 55, Appl
13	141.5	14.6	1161	2 US-08-605-672-55	Sequence 55, Appl
14	141.5	14.6	1161	2 US-08-482-293A-55	Sequence 55, Appl
15	141.5	14.6	1161	2 US-08-943-363-55	Sequence 55, Appl
16	141.5	14.6	1161	3 US-09-193-043-55	Sequence 55, Appl
17	141.5	14.6	1161	4 US-09-688-307A-55	Sequence 55, Appl
18	141.5	14.6	1161	4 US-09-350-259-55	Sequence 55, Appl
19	138.5	14.3	1155	1 US-08-286-889-46	Sequence 46, Appl
20	138.5	14.3	1155	1 US-08-485-618-46	Sequence 46, Appl
21	138.5	14.3	1155	2 US-08-362-652-46	Sequence 46, Appl
22	138.5	14.3	1155	2 US-08-605-672-46	Sequence 46, Appl
23	138.5	14.3	1155	2 US-08-482-293A-46	Sequence 46, Appl
24	138.5	14.3	1155	2 US-08-943-363-46	Sequence 46, Appl
25	138.5	14.3	1155	3 US-09-193-043-46	Sequence 46, Appl
26	138.5	14.3	1155	4 US-09-688-307A-46	Sequence 46, Appl
27	138.5	14.3	1155	4 US-09-350-259-46	Sequence 46, Appl

28	138.5	14.3	1161	1 US-08-485-618-53	Sequence 53, Appl
29	138.5	14.3	1161	1 US-08-362-652-53	Sequence 53, Appl
30	138.5	14.3	1161	2 US-08-605-672-53	Sequence 53, Appl
31	138.5	14.3	1161	2 US-08-482-293A-53	Sequence 53, Appl
32	138.5	14.3	1161	2 US-08-943-363-53	Sequence 53, Appl
33	138.5	14.3	1161	3 US-09-193-043-53	Sequence 53, Appl
34	138.5	14.3	1161	4 US-09-688-307A-53	Sequence 53, Appl
35	138.5	14.3	1161	4 US-09-350-259-53	Sequence 53, Appl
36	136.5	14.1	216	4 US-09-795-872-5	Sequence 5, Appl
37	132.5	13.7	1152	2 US-08-476-062A-43	Sequence 43, Appl
38	132.5	13.7	1152	5 PCT-US96-01314-43	Sequence 43, Appl
39	132.5	13.7	1152	6 5424399-2	Patent No. 5424399
40	132.5	13.7	1152	6 5424399-2	Patent No. 5424399
41	132.5	13.7	1153	1 US-08-173-497-3	Sequence 3, Appl
42	132.5	13.7	1153	1 US-08-286-889-3	Sequence 3, Appl
43	132.5	13.7	1153	1 US-08-485-618-3	Sequence 3, Appl
44	132.5	13.7	1153	1 US-08-362-652-3	Sequence 3, Appl
45	132.5	13.7	1153	2 US-08-605-672-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
PCT-US95-04439-1
Sequence 1, Application PC/TUS9504439
GENERAL INFORMATION:
APPLICANT: The Upjohn Company
APPLICANT: Heinrichson, Robert L.
APPLICANT: Anderson, Donald C.
APPLICANT: Tomlich, Che-Shen C.
APPLICANT: Tombanks, Michael B.
APPLICANT: Bajt, Mary L.
TITLE OF INVENTION: MAC-1-I-DOMAIN PROTEIN USEFUL IN
BLOCKING ADHESION AND MIGRATION OF NEUTROPHILS
TITLE OF INVENTION: BLOCKING ADHESION AND MIGRATION OF NEUTROPHILS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESSES:
ADDRESSER: The Upjohn Company, Intellectual Property Law
STREET: 301 Henrietta
CITY: Kalamazoo
STATE: MI
COUNTRY: USA
ZIP: 49001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Gateway 2000 P5-90
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04439
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Darnley, James D., Jr.
REGISTRATION NUMBER: 33,673
REFERENCE/DOCKET NUMBER: 4767.P CN1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 616/385-5210
TELEFAX: 616/385-6897
TELEX: 224401
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 435 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-04439-1

Query Match 14.7% ; Score 142.5 ; DB 5 ; Length 435 ;
Best Local Similarity 26.9% ; Pred. No. 8.4e-08 ;
Matches 54 ; Conservative 41 ; Mismatches 67 ; Indels 39 ; Gaps 11 ;

Db 257 PLEYSVDIPADKX---GIIRYALGVGDAPQEPFALKEINTIGSAPPODHVFKVGN-FAA 312

QY 168 LOGIHSILKK 178

Db 313 LRSIORQLOEK 323

RESULT 4

US-08-362-652-37

Sequence 37, Application US/08362652

Patent No. 5766850

GENERAL INFORMATION:

APPLICANT: Gallatin, W. Michael

APPLICANT: Van der Vlieten, Monica

TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit

NUMBER OF SEQUENCES: 93

CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

STREET: 233 South Wacker Drive, 6300 Sear Tower

CITY: Chicago

STATE: Illinois

COUNTRY: United States

ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/362,652

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/173,497

FILING DATE: 23-DEC-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/286,889

FILING DATE: 5-AUG-1994

ATTORNEY/AGENT INFORMATION:

NAME: Williams Jr., Joseph A.

REGISTRATION NUMBER: 38,659

REFERENCE/DOCKET NUMBER: 27866/32391

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-474-6300

TELEFAX: 312-474-0448

TELEX: 25-3856

INFORMATION FOR SEO ID NO: 37:

SEQUENCE CHARACTERISTICS:

LENGTH: 1151 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-362-652-37

Query Match 14.6%; Score 141.5; DB 1; Length 1151;

Best Local Similarity 27.7%; Pred. No. 4.5e-07;

Matches 53; Conservative 38; Mismatches 75; Indels 25; Gaps 10;

QY 4 DLVFIIDKGSV-LHHMNEIYFVEQLAKFISPOLRMSFIVSTGTLMLKTE----- 57

Db 142 DIAVFLIDSGSGSINQORDFAQMKDFVKALMGEPASTSTLFSIMQYSNLIKHTFTTEPKNII 201

QY 58 DREQIQGLLELQKVLPGSDTYMHGEPERASQIYYENRGVRYTA-SVIALTDGELHED 116

Db 202 DPQSLVDPIVQLQ-----GLTYTATGIRYMEELFHSKSGSRKSAKKILIVITDQKIRN 256

QY 117 LFFYSE--BEANRSRLGAIIVYCVGYK-FNE-TOLARI-----ADSKDHVFPVNDGFOA 167

Db 257 PLEYSVDIPADKX---GIIRYALGVGDAPQEPFALKEINTIGSAPPODHVFKVGN-FAA 312

QY 168 LOGIHSILKK 178

Db 313 LRSIORQLOEK 323

RESULT 5

US-08-605-672-37

Sequence 37, Application US/08605672

Patent No. 5817515

GENERAL INFORMATION:

APPLICANT: Gallatin, W. Michael

APPLICANT: Van der Vlieten, Monica

TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit

NUMBER OF SEQUENCES: 103

CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

STREET: 233 South Wacker Drive, 6300 Sear Tower

CITY: Chicago

STATE: Illinois

COUNTRY: United States

ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/605,672

FILING DATE:

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/173,497

FILING DATE: 23-DEC-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/286,889

FILING DATE: 5-AUG-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/362,652

FILING DATE: 21-DEC-1994

ATTORNEY/AGENT INFORMATION:

NAME: Williams Jr., Joseph A.

REGISTRATION NUMBER: 38,659

REFERENCE/DOCKET NUMBER: 27866/32684

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-474-6300

TELEFAX: 312-474-0448

TELEX: 25-3856

INFORMATION FOR SEO ID NO: 37:

SEQUENCE CHARACTERISTICS:

LENGTH: 1151 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-605-672-37

Query Match 14.6%; Score 141.5; DB 2; Length 1151;

Best Local Similarity 27.7%; Pred. No. 4.5e-07;

Matches 53; Conservative 38; Mismatches 75; Indels 25; Gaps 10;

QY 4 DLVFIIDKGSV-LHHMNEIYFVEQLAKFISPOLRMSFIVSTGTLMLKTE----- 57

Db 142 DIAVFLIDSGSGSINQORDFAQMKDFVKALMGEPASTSTLFSIMQYSNLIKHTFTTEPKNII 201

QY 58 DREQIQGLLELQKVLPGSDTYMHGEPERASQIYYENRGVRYTA-SVIALTDGELHED 116

Db 202 DPQSLVDPIVQLQ-----GLTYTATGIRYMEELFHSKSGSRKSAKKILIVITDQKIRN 256

QY 117 LFFYSE--BEANRSRLGAIIVYCVGYK-FNE-TOLARI-----ADSKDHVFPVNDGFOA 167

Db 257 PLEYSVDIPADKX---GIIRYALGVGDAPQEPFALKEINTIGSAPPODHVFKVGN-FAA 312

QY 168 LOGIHSILKK 178

Db 313 LRSIORQLOEK 323

```
RESULT 6
US-08-482-293A-37
; Sequence 37, Application US/08482293A
; Patent No. 5831029
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vlieten, Monica
; TITLE OF INVENTION: No. 5831029e1 Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,293A
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1151 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-482-293A-37

Query Match 14.6%; Score 141.5; DB 2; Length 1151;
Best Local Similarity 27.7%; Pred. NO. 4.5e-07;
Matches 53; Conservative 38; Mismatches 75; Indels 25; Gaps 10;

QY 4 DLYFILDKSGSV-LHHNNEIYFVEQLAHKEISPOLRMSFIVSTRTGLMKLTE----- 57
DB 142 DIALFDSSGSSINQRDPQKMDVFKALMGFFASTSTLFSLMQYSNLIKTHFTFEPKNIL 201
QY 58 DREQIRGLBELQKVLPGGDTYMHGFERASEQIYYENRGYRTA-SVITALTGDELHED 116
DB 202 DPGSLVDPIVQLQ-----GLTYTATGIRTYMEELFHSKNSRSKAKILLVITDGQKXRD 256
QY 117 LFFYSE--REANRSDIGAIVYCVKVD-FNE-TQLARI-----ADSKDHVFPVNDGFOA 167
DB 257 PLEYSVDVIPADKA---GIIRYALGVDAFOEPTALKELNTIGSAPPQDHVFKVGN-FAA 312
QY 168 LOGIHSILKK 178
DB 313 LRSTORQLOEK 323

RESULT 7
US-08-943-363-37
```

```
; Sequence 37, Application US/08943363
; Patent No. 5837478
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vlieten, Monica
; TITLE OF INVENTION: No. 5837478e1 Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,363
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1151 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-943-363-37

Query Match 14.6%; Score 141.5; DB 2; Length 1151;
Best Local Similarity 27.7%; Pred. NO. 4.5e-07;
Matches 53; Conservative 38; Mismatches 75; Indels 25; Gaps 10;

QY 4 DLYFILDKSGSV-LHHNNEIYFVEQLAHKEISPOLRMSFIVSTRTGLMKLTE----- 57
DB 142 DIALFDSSGSSINQRDPQKMDVFKALMGFFASTSTLFSLMQYSNLIKTHFTFEPKNIL 201
QY 58 DREQIRGLBELQKVLPGGDTYMHGFERASEQIYYENRGYRTA-SVITALTGDELHED 116
DB 202 DPGSLVDPIVQLQ-----GLTYTATGIRTYMEELFHSKNSRSKAKILLVITDGQKXRD 256
QY 117 LFFYSE--REANRSDIGAIVYCVKVD-FNE-TQLARI-----ADSKDHVFPVNDGFOA 167
DB 257 PLEYSVDVIPADKA---GIIRYALGVDAFOEPTALKELNTIGSAPPQDHVFKVGN-FAA 312
QY 168 LOGIHSILKK 178
DB 313 LRSTORQLOEK 323

RESULT 8
US-09-193-043-37
; Sequence 37, Application US/09193043
; Patent No. 6251395
```

GENERAL INFORMATION:
APPLICANT: Gallatin, Michael W.
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 6251395e1 Human 2
FILE REFERENCE: 27866/35004
CURRENT APPLICATION NUMBER: US/09/193,043
PRIOR FILING DATE: 1998-11-16
EARLIER APPLICATION NUMBER: 08/173,497
PRIOR FILING DATE: 1993-12-23
EARLIER APPLICATION NUMBER: 08/286,889
PRIOR FILING DATE: 1994-08-05
EARLIER APPLICATION NUMBER: 08/362,652
PRIOR FILING DATE: 1994-12-21
EARLIER APPLICATION NUMBER: 08/943,363
NUMBER OF SEQ ID NOS: 114
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 37
LENGTH: 1151
TYPE: PRT
ORGANISM: Rattus rattus
US-09-193-043-37

Query Match 14.6%; Score 141.5; DB 3; Length 1151;
Best Local Similarity 27.7%; Pred. No. 4.5e-07;
Matches 53; Conservative 38; Mismatches 75; Indels 25; Gaps 10;
Qy 4 DLYFLDKSGSV-LHNMNEIYFVEQLAHKFIQPOLRMSPIVSTRTGTLMLKTE----- 57
Db 142 DIAFLDSSGSSINORDPAQKMDFYKALMGEPASTSTFLSMQYSNLKHTFTFEPKNIL 201
Qy 58 DREGIROGLBELQVLPFGDTYMEGPERASEQIYYENRQGYTA-SVIALTDGELHED 116
Db 202 DPGSLDPIVQLQ-----GLTYATGIRTWBELFHSKNGSRKSAKKILLVITDQCKRD 256
Qy 117 LFPYSE--REANRSDGAIYVCYKVD-FNE-TQLARI-----ADSKDHVPVNDGFOA 167
Db 257 PLEYSDVIPADAKA---GIIRYALGVGDAPQEBTALKEINTIGSAPPDHVFYKGN-FAA 312
Qy 168 LOGIHSILKK 178
Db 313 LRSIORQLQEK 323

RESULT 9
US-09-688-307A-37
Sequence 37, Application US/09688307A
Patent No. 6432404
GENERAL INFORMATION:
APPLICANT: Gallatin, Michael W.
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 6432404e1 Human Beta-2
FILE REFERENCE: 27866/36646
CURRENT APPLICATION NUMBER: US/09/688,307A
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 09/193,043
PRIOR FILING DATE: 1998-11-16
PRIOR APPLICATION NUMBER: 08/605,672
PRIOR FILING DATE: 1996-02-22
PRIOR APPLICATION NUMBER: 08/173,497
PRIOR FILING DATE: 1993-12-23
PRIOR APPLICATION NUMBER: 08/286,889
PRIOR FILING DATE: 1994-08-05
PRIOR APPLICATION NUMBER: 08/362,652
PRIOR FILING DATE: 1994-12-21
PRIOR APPLICATION NUMBER: 08/943,363
NUMBER OF SEQ ID NOS: 114
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 37
LENGTH: 1151
TYPE: PRT
ORGANISM: Rattus rattus

FEATURE:
NAME/KEY: misc_feature
LOCATION: 361
OTHER INFORMATION: Xaa = any or unknown amino acid
NAME/KEY: misc_feature
LOCATION: 464
OTHER INFORMATION: Xaa = any or unknown amino acid
NAME/KEY: misc_feature
LOCATION: 486
OTHER INFORMATION: Xaa = any or unknown amino acid
NAME/KEY: misc_feature
LOCATION: 506
OTHER INFORMATION: Xaa = any or unknown amino acid
NAME/KEY: misc_feature
LOCATION: 1117
OTHER INFORMATION: Xaa = any or unknown amino acid
NAME/KEY: misc_feature
LOCATION: 1118
OTHER INFORMATION: Xaa = any or unknown amino acid
US-09-688-307A-37

Query Match 14.6%; Score 141.5; DB 4; Length 1151;
Best Local Similarity 27.7%; Pred. No. 4.5e-07;
Matches 53; Conservative 38; Mismatches 75; Indels 25; Gaps 10;
Qy 4 DLYFLDKSGSV-LHNMNEIYFVEQLAHKFIQPOLRMSPIVSTRTGTLMLKTE----- 57
Db 142 DIAFLDSSGSSINORDPAQKMDFYKALMGEPASTSTFLSMQYSNLKHTFTFEPKNIL 201
Qy 58 DREGIROGLBELQVLPFGDTYMEGPERASEQIYYENRQGYTA-SVIALTDGELHED 116
Db 202 DPGSLDPIVQLQ-----GLTYATGIRTWBELFHSKNGSRKSAKKILLVITDQCKRD 256
Qy 117 LFPYSE--REANRSDGAIYVCYKVD-FNE-TQLARI-----ADSKDHVPVNDGFOA 167
Db 257 PLEYSDVIPADAKA---GIIRYALGVGDAPQEBTALKEINTIGSAPPDHVFYKGN-FAA 312
Qy 168 LOGIHSILKK 178
Db 313 LRSIORQLQEK 323

RESULT 10
US-09-350-259-37
Sequence 37, Application US/09350259
Patent No. 6620915
GENERAL INFORMATION:
APPLICANT: Gallatin, Michael W.
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 6620915e1 Human 2
FILE REFERENCE: 27866/35004
CURRENT APPLICATION NUMBER: US/09/350,259
PRIOR FILING DATE: 1999-07-08
EARLIER APPLICATION NUMBER: 09/193,043
PRIOR FILING DATE: 1998-11-16
EARLIER APPLICATION NUMBER: 08/173,497
PRIOR FILING DATE: 1993-12-23
EARLIER APPLICATION NUMBER: 08/286,889
PRIOR FILING DATE: 1994-08-05
EARLIER APPLICATION NUMBER: 08/362,652
PRIOR FILING DATE: 1994-12-21
EARLIER APPLICATION NUMBER: 08/943,363
NUMBER OF SEQ ID NOS: 114
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 37
LENGTH: 1151
TYPE: PRT
ORGANISM: Rattus rattus
US-09-350-259-37

Query Match 14.6%; Score 141.5; DB 4; Length 1151;
Best Local Similarity 27.7%; Pred. No. 4.5e-07;

Db 212 DPOSVDPIVQLO-----GLTYTATGIRTYWELFHSKNGSRKSAKILLVITDGQKRD 266
Qy 117 LFFYSR--REANRSDIGAIVYCVGVD-FNE-TQIARI-----ADSKDHVPVNDGFOA 167
Db 267 PLEYSVDVIPADKX-----GIIRYAIQVGDARFOEPALKEINLTIGSAPPQDHVFKVGN-FAA 322
Qy 168 LOGIHSILKK 178
Db 323 LRSTQRQLQEK 333

RESULT 13

US-08-605-672-55
Sequence 55, Application US/08605672
Patent No. 5817515
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/605.672
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 1161 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-605-672-55

Query Match 14.6%; Score 141.5; DB 2; Length 1161;
Best Local Similarity 27.7%; Pred. No. 4.5e-07;
Matches 53; Conservative 38; Mismatches 75; Indels 25; Gaps 10;
Qy 4 DLYFIIDKSGSV-LHHNNEIYVVEQLAHKFI SPQLRMSFIVFSTGTTLMKLT-----57
Db 152 DIAFLIDGSGSINQRDPAQWKDFV KALMGEPASTSTLFSIMQYSNLIKTHFTFEFNIL 211
Qy 58 DRQIRIGLELEQVLRGSDTYVHNGEPERASBOIYYENQGYRTA-SVIALALTDGELHED 116
Db 212 DPOSVDPIVQLO-----GLTYTATGIRTYWELFHSKNGSRKSAKILLVITDGQKRD 266

Qy 117 LFFYSR--REANRSDIGAIVYCVGVD-FNE-TQIARI-----ADSKDHVPVNDGFOA 167
Db 267 PLEYSVDVIPADKX-----GIIRYAIQVGDARFOEPALKEINLTIGSAPPQDHVFKVGN-FAA 322
Qy 168 LOGIHSILKK 178
Db 323 LRSTQRQLQEK 333

RESULT 14

US-08-482-293A-55
Sequence 55, Application US/08482293A
Patent No. 5831029
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482.293A
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 1161 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-482-293A-55

Query Match 14.6%; Score 141.5; DB 2; Length 1161;
Best Local Similarity 27.7%; Pred. No. 4.5e-07;
Matches 53; Conservative 38; Mismatches 75; Indels 25; Gaps 10;
Qy 4 DLYFIIDKSGSV-LHHNNEIYVVEQLAHKFI SPQLRMSFIVFSTGTTLMKLT-----57
Db 152 DIAFLIDGSGSINQRDPAQWKDFV KALMGEPASTSTLFSIMQYSNLIKTHFTFEFNIL 211
Qy 58 DRQIRIGLELEQVLRGSDTYVHNGEPERASBOIYYENQGYRTA-SVIALALTDGELHED 116
Db 212 DPOSVDPIVQLO-----GLTYTATGIRTYWELFHSKNGSRKSAKILLVITDGQKRD 266
Qy 117 LFFYSR--REANRSDIGAIVYCVGVD-FNE-TQIARI-----ADSKDHVPVNDGFOA 167

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 20:03:29 ; Search time 59.9593 Seconds
(without alignments)
1195.537 Million cell updates/sec

Title: US-09-970-076-2_COPY_41_227

Perfect score: 970
Sequence: 1 GGFPLYILDKSGSVLHNMN.....LOGIHSILKSCIEILAE 187

Scoring table: BLASTSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1710399 seqs, 383334425 residues

Total number of hits satisfying chosen parameters: 1710399

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubppa/PC7_NEM_PUB.pep.*
3: /cgn2_6/ptodata/1/pubppa/US06_NEM_PUB.pep.*
4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubppa/US07_NEM_PUB.pep.*
6: /cgn2_6/ptodata/1/pubppa/PC7US_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubppa/US08_NEM_PUB.pep.*
8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubppa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubppa/US09_NEM_PUB.pep.*
13: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubppa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubppa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubppa/US10_NEM_PUB.pep.*
19: /cgn2_6/ptodata/1/pubppa/US11_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubppa/US11_NEM_PUB.pep.*
21: /cgn2_6/ptodata/1/pubppa/US60_NEM_PUB.pep.*
22: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	970	100.0	328 14 US-10-038-307-26	Sequence 26, Appl
2	970	100.0	328 14 US-10-201-292-26	Sequence 26, Appl
3	970	100.0	333 10 US-09-796-753-12	Sequence 12, Appl
4	970	100.0	333 14 US-10-038-307-2	Sequence 2, Appl
5	970	100.0	333 14 US-10-201-292-2	Sequence 2, Appl
6	970	100.0	342 14 US-10-038-307-22	Sequence 22, Appl
7	970	100.0	342 14 US-10-201-292-22	Sequence 22, Appl
8	970	100.0	345 14 US-10-038-307-24	Sequence 24, Appl
9	970	100.0	345 14 US-10-201-292-24	Sequence 24, Appl
10	970	100.0	403 11 US-09-833-245-621	Sequence 621, Appl
11	970	100.0	460 14 US-10-201-292-28	Sequence 28, Appl

12	970	100.0	479 14 US-10-201-292-32	Sequence 32, Appl
13	970	100.0	504 14 US-10-201-292-34	Sequence 34, Appl
14	970	100.0	529 14 US-10-201-292-36	Sequence 36, Appl
15	970	100.0	551 14 US-10-038-307-18	Sequence 18, Appl
16	970	100.0	551 14 US-10-201-292-18	Sequence 18, Appl
17	970	100.0	564 10 US-09-818-715-187	Sequence 187, Appl
18	970	100.0	564 10 US-09-818-715-232	Sequence 232, Appl
19	970	100.0	564 14 US-10-038-307-20	Sequence 20, Appl
20	970	100.0	564 14 US-10-201-292-199	Sequence 199, Appl
21	970	100.0	564 14 US-10-301-822-12	Sequence 1823, Ap
22	970	100.0	564 16 US-10-408-7654-1823	Sequence 187, Ap
23	970	100.0	564 16 US-10-474-794-187	Sequence 232, App
24	970	100.0	564 16 US-10-474-794-232	Sequence 620, App
25	965	99.5	403 11 US-09-833-245-620	Sequence 194, App
26	961	99.1	562 10 US-09-818-715-194	Sequence 301, App
27	961	99.1	562 10 US-09-818-715-301	Sequence 194, App
28	961	99.1	562 16 US-10-474-794-194	Sequence 301, App
29	961	99.1	562 16 US-10-474-794-301	Sequence 301, App
30	960	99.0	460 14 US-10-201-292-30	Sequence 12, Appl
31	949	97.8	534 14 US-10-038-307-12	Sequence 12, Appl
32	949	97.8	534 14 US-10-201-292-12	Sequence 10, Appl
33	949	97.8	543 14 US-10-038-307-10	Sequence 10, Appl
34	949	97.8	543 14 US-10-038-307-14	Sequence 14, Appl
35	949	97.8	543 14 US-10-038-307-16	Sequence 16, Appl
36	949	97.8	543 14 US-10-201-292-16	Sequence 16, Appl
37	949	97.8	543 14 US-10-201-292-14	Sequence 14, Appl
38	949	97.8	538 13 US-10-047-542-99	Sequence 99, Appl
39	900	92.8	538 13 US-10-047-542-99	Sequence 99, Appl
40	550	56.7	488 14 US-10-038-307-6	Sequence 6, Appl
41	550	56.7	488 14 US-10-201-292-6	Sequence 6, Appl
42	550	56.7	488 14 US-10-201-292-6	Sequence 16, Appl
43	550	56.7	488 14 US-10-368-087-16	Sequence 2639, Ap
44	550	56.7	488 15 US-10-104-047-2639	Sequence 54, Appl
45	548	56.5	487 10 US-09-796-753-54	

ALIGNMENTS

RESULT 1
US-10-038-307-26
; Sequence 26, Application US/10038307
; Publication No. US20030134786A1
; GENERAL INFORMATION:
; APPLICANT: James B. ROTTMAN
; APPLICANT: Theresa L. O'KEEFE
; APPLICANT: Engin OZKAYNAK
; APPLICANT: Judith J. HEALEY
; TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
; FILE REFERENCE: 7853-253-999
; CURRENT APPLICATION NUMBER: US/10/038,307
; CURRENT FILING DATE: 2002-06-28
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-038-307-26
Query Match 100.0%; Score 970; DB 14; Length 328;
Best Local Similarity 100.0%; Pred. No. 6e-93;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGFPLYILDKSGSVLHNMNBIYFVQLAHKPIPSPOLRMSFIYFSTRGTTMLKLTEDRE 60
DB 41 GGFPLYILDKSGSVLHNMNBIYFVQLAHKPIPSPOLRMSFIYFSTRGTTMLKLTEDRE 100
QY 61 QIRGAEELQKVPKGGDTYHGBEPERASEQIYYENRGYRTASYITLITDGELEDDLFY 120
DB 101 QIRGAEELQKVPKGGDTYHGBEPERASEQIYYENRGYRTASYITLITDGELEDDLFY 160
QY 121 SREANRSDLGAIVYCVGVKDFNETOLARIADSKDHFVNDGFOALOGIHSILKSC 180

```
Db      161 SEREANSRDLGAIYCVGVKDPNETOLARIADSKDHVPVNDGFOALOGIIHSILKKSC 220
Qy      181 IETIAAE 187
Db      221 IETIAAE 227

RESULT 2
US-10-201-292-26
; Sequence 26, Application US/10201292
; Publication No. US20030144193A1
; GENERAL INFORMATION:
; APPLICANT: James B. ROTTMAN
; APPLICANT: Inereba L. O'KEEFE
; APPLICANT: Engin OZKAYNAK
; APPLICANT: Judith J. HEALEY
; TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
; FILE REFERENCE: 7853-253-999
; CURRENT APPLICATION NUMBER: US/10/201,292
; CURRENT FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-201-292-26

Query Match      100.0%; Score 970; DB 14; Length 328;
Best Local Similarity 100.0%; Pred. No. 6e-93;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGFDFLYFLIDKSGSVLHNNNEIYFVEQLAHKFISPOLRMSFIVFSTRGTTLMKLTEDRE 60
Db      41 GGFDFLYFLIDKSGSVLHNNNEIYFVEQLAHKFISPOLRMSFIVFSTRGTTLMKLTEDRE 100
Qy      61 QIRQGLEELQKVLPGGDTYTHHEGFERASEQIYYENRGQYRTASVITLTDGELHEDLFFY 120
Db      101 QIRQGLEELQKVLPGGDTYTHHEGFERASEQIYYENRGQYRTASVITLTDGELHEDLFFY 160
Qy      121 SEREANSRDLGAIYCVGVKDPNETOLARIADSKDHVPVNDGFOALOGIIHSILKKSC 180
Db      161 SEREANSRDLGAIYCVGVKDPNETOLARIADSKDHVPVNDGFOALOGIIHSILKKSC 220
Qy      181 IETIAAE 187
Db      221 IETIAAE 227

RESULT 3
US-09-796-753-12
; Sequence 12, Application US/09796753
; Publication No. US20030027998A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-227-999
; CURRENT APPLICATION NUMBER: US/09/796,753
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 09/183,175
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 09/223,094
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/223,546
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/224,246
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/259,388
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 60/122,458
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 09/312,359
```

```
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/336,536
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: 09/342,687
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 09/345,464
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: 09/365,164
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 09/399,723
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 09/409,634
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 09/471,179
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 09/474,071
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/474,072
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/514,010
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: 09/516,745
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/572,002
; PRIOR FILING DATE: 2000-05-14
; PRIOR APPLICATION NUMBER: 09/597,993
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 09/599,596
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 09/630,334
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: 09/606,565
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/606,317
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/665,666
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 09/677,751
; PRIOR FILING DATE: 2000-09-30
; NUMBER OF SEQ ID NOS: 162
; SEQ ID NO 12
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-796-753-12

Query Match      100.0%; Score 970; DB 10; Length 333;
Best Local Similarity 100.0%; Pred. No. 6.2e-93;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGFDFLYFLIDKSGSVLHNNNEIYFVEQLAHKFISPOLRMSFIVFSTRGTTLMKLTEDRE 60
Db      41 GGFDFLYFLIDKSGSVLHNNNEIYFVEQLAHKFISPOLRMSFIVFSTRGTTLMKLTEDRE 100
Qy      61 QIRQGLEELQKVLPGGDTYTHHEGFERASEQIYYENRGQYRTASVITLTDGELHEDLFFY 120
Db      101 QIRQGLEELQKVLPGGDTYTHHEGFERASEQIYYENRGQYRTASVITLTDGELHEDLFFY 160
Qy      121 SEREANSRDLGAIYCVGVKDPNETOLARIADSKDHVPVNDGFOALOGIIHSILKKSC 180
Db      161 SEREANSRDLGAIYCVGVKDPNETOLARIADSKDHVPVNDGFOALOGIIHSILKKSC 220
Qy      181 IETIAAE 187
Db      221 IETIAAE 227

RESULT 4
US-10-038-307-2
; Sequence 2, Application US/10038307
; Publication No. US20030134786A1
; GENERAL INFORMATION:
; APPLICANT: James B. ROTTMAN
```

```
/ APPLICANT: Theresa L. O'KEEFE
/ APPLICANT: Engin OZKAYNAK
/ APPLICANT: Judith J. HEALEY
/ TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
/ FILE REFERENCE: 7853-253-999
/ CURRENT APPLICATION NUMBER: US/10/038,307
/ CURRENT FILING DATE: 2002-06-28
/ NUMBER OF SEQ ID NOS: 26
/ SOFTWARE: PaSeq for Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 333
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-038-307-2
```

```
Query Match      100.0%; Score 970; DB 14; Length 333;
Best Local Similarity 100.0%; Pred. No. 6,2e-93;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GGFDFYFIIDKSGSVLHNMNEIYFVEQLAHKFIISPOLRMSFIVFSTRGTTLMKLTEDRE 60
DB 41 GGFDFYFIIDKSGSVLHNMNEIYFVEQLAHKFIISPOLRMSFIVFSTRGTTLMKLTEDRE 100
QY 61 QIRQGLEELQKVLPGGDTYHHEGFERASEQIYYENRGYRTASVITALTGELHEDLFFY 120
DB 101 QIRQGLEELQKVLPGGDTYHHEGFERASEQIYYENRGYRTASVITALTGELHEDLFFY 160
QY 121 SEREANSRDLGAIIVCVGVKDFNETQIARIADSKDHVFPVNDGFQALOGIHSILKKSC 180
DB 161 SEREANSRDLGAIIVCVGVKDFNETQIARIADSKDHVFPVNDGFQALOGIHSILKKSC 220
QY 181 IEILAAE 187
DB 221 IEILAAE 227
```

```
RESULT 5
US-10-201-292-2
/ Sequence 22, Application US/10201292
/ Publication No. US20030144193A1
/ GENERAL INFORMATION:
/ APPLICANT: James B. ROTTMAN
/ APPLICANT: Theresa L. O'KEEFE
/ APPLICANT: Engin OZKAYNAK
/ APPLICANT: Judith J. HEALEY
/ TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
/ FILE REFERENCE: 7853-253-999
/ CURRENT APPLICATION NUMBER: US/10/201,292
/ CURRENT FILING DATE: 2003-02-14
/ NUMBER OF SEQ ID NOS: 36
/ SOFTWARE: PaSeq for Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 333
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-201-292-2
```

```
Query Match      100.0%; Score 970; DB 14; Length 333;
Best Local Similarity 100.0%; Pred. No. 6,2e-93;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GGFDFYFIIDKSGSVLHNMNEIYFVEQLAHKFIISPOLRMSFIVFSTRGTTLMKLTEDRE 60
DB 41 GGFDFYFIIDKSGSVLHNMNEIYFVEQLAHKFIISPOLRMSFIVFSTRGTTLMKLTEDRE 100
QY 61 QIRQGLEELQKVLPGGDTYHHEGFERASEQIYYENRGYRTASVITALTGELHEDLFFY 120
DB 101 QIRQGLEELQKVLPGGDTYHHEGFERASEQIYYENRGYRTASVITALTGELHEDLFFY 160
QY 121 SEREANSRDLGAIIVCVGVKDFNETQIARIADSKDHVFPVNDGFQALOGIHSILKKSC 180
DB 161 SEREANSRDLGAIIVCVGVKDFNETQIARIADSKDHVFPVNDGFQALOGIHSILKKSC 220
```

```
QY 181 IEILAAE 187
DB 221 IEILAAE 227
```

```
RESULT 6
US-10-038-307-22
/ Sequence 22, Application US/10038307
/ Publication No. US20030134786A1
/ GENERAL INFORMATION:
/ APPLICANT: James B. ROTTMAN
/ APPLICANT: Theresa L. O'KEEFE
/ APPLICANT: Engin OZKAYNAK
/ APPLICANT: Judith J. HEALEY
/ TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
/ FILE REFERENCE: 7853-253-999
/ CURRENT APPLICATION NUMBER: US/10/038,307
/ CURRENT FILING DATE: 2002-06-28
/ NUMBER OF SEQ ID NOS: 26
/ SOFTWARE: PaSeq for Windows Version 4.0
/ SEQ ID NO 22
/ LENGTH: 342
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-038-307-22
```

```
Query Match      100.0%; Score 970; DB 14; Length 342;
Best Local Similarity 100.0%; Pred. No. 6,4e-93;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GGFDFYFIIDKSGSVLHNMNEIYFVEQLAHKFIISPOLRMSFIVFSTRGTTLMKLTEDRE 60
DB 50 GGFDFYFIIDKSGSVLHNMNEIYFVEQLAHKFIISPOLRMSFIVFSTRGTTLMKLTEDRE 109
QY 61 QIRQGLEELQKVLPGGDTYHHEGFERASEQIYYENRGYRTASVITALTGELHEDLFFY 120
DB 110 QIRQGLEELQKVLPGGDTYHHEGFERASEQIYYENRGYRTASVITALTGELHEDLFFY 169
QY 121 SEREANSRDLGAIIVCVGVKDFNETQIARIADSKDHVFPVNDGFQALOGIHSILKKSC 180
DB 170 SEREANSRDLGAIIVCVGVKDFNETQIARIADSKDHVFPVNDGFQALOGIHSILKKSC 229
QY 181 IEILAAE 187
DB 230 IEILAAE 236
```

```
RESULT 7
US-10-201-292-22
/ Sequence 22, Application US/10201292
/ Publication No. US20030144193A1
/ GENERAL INFORMATION:
/ APPLICANT: James B. ROTTMAN
/ APPLICANT: Theresa L. O'KEEFE
/ APPLICANT: Engin OZKAYNAK
/ APPLICANT: Judith J. HEALEY
/ TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
/ FILE REFERENCE: 7853-253-999
/ CURRENT APPLICATION NUMBER: US/10/201,292
/ CURRENT FILING DATE: 2003-02-14
/ NUMBER OF SEQ ID NOS: 36
/ SOFTWARE: PaSeq for Windows Version 4.0
/ SEQ ID NO 22
/ LENGTH: 342
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-201-292-22
```

```
Query Match      100.0%; Score 970; DB 14; Length 342;
Best Local Similarity 100.0%; Pred. No. 6,4e-93;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GGFDFYFIIDKSGSVLHNMNEIYFVEQLAHKFIISPOLRMSFIVFSTRGTTLMKLTEDRE 60
```

```

Db      50  GGFDDYFLIDKSGSVLHHMNEIYYVEQLAHKFISPOLMSFIVSTRTGLTMKLTEDRE 109
QY      61  QIROGLEELQKVLPGGDTYYMEGFERASQIYYENROGRTASVIALTDGELHEDLFFY 120
Db      110  QIROGLEELQKVLPGGDTYYMEGFERASQIYYENROGRTASVIALTDGELHEDLFFY 169
QY      121  SERENRNRSDGAIYYCVGVKDFNETOLARIADSKDHFVPVNDGFQALOGIITHSLTKKSC 180
Db      170  SERENRNRSDGAIYYCVGVKDFNETOLARIADSKDHFVPVNDGFQALOGIITHSLTKKSC 229
QY      181  IEIILAAE 187
      |||||
Db      230  IEIILAAE 236

RESULT 8
US-10-038-307-24
; Sequence 24, Application US/10038307
; Publication No. US20030134786A1
; GENERAL INFORMATION:
; APPLICANT: James B. ROTTMAN
; APPLICANT: Theresa L. O'KEEFE
; APPLICANT: Engin OZKANNAK
; APPLICANT: Judith J. HEALEY
; TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
; FILE REFERENCE: 7853-253-899
; CURRENT APPLICATION NUMBER: US/10/038,307
; CURRENT FILING DATE: 2002-06-28
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-038-307-24

Query Match      100.0%; SCORE 970; DB 14; Length 345;
Best Local Similarity 100.0%; Freq. No. 6,5e-93;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY      1  GGFDDYFLIDKSGSVLHHMNEIYYVEQLAHKFISPOLMSFIVSTRTGLTMKLTEDRE 60
Db      41  GGFDDYFLIDKSGSVLHHMNEIYYVEQLAHKFISPOLMSFIVSTRTGLTMKLTEDRE 100
QY      61  QIROGLEELQKVLPGGDTYYMEGFERASQIYYENROGRTASVIALTDGELHEDLFFY 120
Db      101  QIROGLEELQKVLPGGDTYYMEGFERASQIYYENROGRTASVIALTDGELHEDLFFY 160
QY      121  SERENRNRSDGAIYYCVGVKDFNETOLARIADSKDHFVPVNDGFQALOGIITHSLTKKSC 180
Db      161  SERENRNRSDGAIYYCVGVKDFNETOLARIADSKDHFVPVNDGFQALOGIITHSLTKKSC 220
QY      181  IEIILAAE 187
      |||||
Db      221  IEIILAAE 227

RESULT 9
US-10-201-292-24
; Sequence 24, Application US/10201292
; Publication No. US20030144193A1
; GENERAL INFORMATION:
; APPLICANT: James B. ROTTMAN
; APPLICANT: Theresa L. O'KEEFE
; APPLICANT: Engin OZKANNAK
; APPLICANT: Judith J. HEALEY
; TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
; FILE REFERENCE: 7853-253-999
; CURRENT APPLICATION NUMBER: US/10/201,292
; CURRENT FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

? SEQ ID NO 24
? LENGTH: 345
? TYPE: PRT
? ORGANISM: Homo sapiens
US-10-201-292-24

Query Match
Best Local Similarity 100.0%; Score 970; DB 14; Length 345;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFPLYFLIDKSGSVLHHMNEIYFVEQLAHKFI SPQLRMSFIYFSTRGTTLMKLTEDRE 60
Db 41 GGFPLYFLIDKSGSVLHHMNEIYFVEQLAHKFI SPQLRMSFIYFSTRGTTLMKLTEDRE 100

QY 61 QIRGSELQVLPFGDTYMHGEPERASEQIYYENRGYRTASVIALTDGLHEDLFFY 120
Db 101 QIRGSELQVLPFGDTYMHGEPERASEQIYYENRGYRTASVIALTDGLHEDLFFY 160

QY 121 SERANSRDGLAIVYCVGVDPFNETQLARIADSKDHVPVNDGFQALQGIHSLKKSC 180
Db 161 SERANSRDGLAIVYCVGVDPFNETQLARIADSKDHVPVNDGFQALQGIHSLKKSC 220

QY 181 IEIILAE 187
Db 221 IEIILAE 227

RESULT 10
US-09-833-245-621
? Sequence 621, Application US/09833245
? Publication No. US20040010134A1
? GENERAL INFORMATION:
? APPLICANT: Human Genome Sciences, Inc.
? TITLE OF INVENTION: Albumin Fusion Proteins
? FILE REFERENCE: PFS46PCT
? CURRENT APPLICATION NUMBER: US/09/833,245
? CURRENT FILING DATE: 2001-04-12
? PRIOR APPLICATION NUMBER: 60/229, 358
? PRIOR FILING DATE: 2000-04-12
? PRIOR APPLICATION NUMBER: 60/256, 931
? PRIOR FILING DATE: 2000-12-21
? PRIOR APPLICATION NUMBER: 60/199, 384
? PRIOR FILING DATE: 2000-04-25
? NUMBER OF SEQ ID NOS: 2267
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 621
? LENGTH: 403
? TYPE: PRT
? ORGANISM: Homo sapiens
US-09-833-245-621

Query Match
Best Local Similarity 100.0%; Score 970; DB 11; Length 403;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFPLYFLIDKSGSVLHHMNEIYFVEQLAHKFI SPQLRMSFIYFSTRGTTLMKLTEDRE 60
Db 41 GGFPLYFLIDKSGSVLHHMNEIYFVEQLAHKFI SPQLRMSFIYFSTRGTTLMKLTEDRE 100

QY 61 QIRGSELQVLPFGDTYMHGEPERASEQIYYENRGYRTASVIALTDGLHEDLFFY 120
Db 101 QIRGSELQVLPFGDTYMHGEPERASEQIYYENRGYRTASVIALTDGLHEDLFFY 160

QY 121 SERANSRDGLAIVYCVGVDPFNETQLARIADSKDHVPVNDGFQALQGIHSLKKSC 180
Db 161 SERANSRDGLAIVYCVGVDPFNETQLARIADSKDHVPVNDGFQALQGIHSLKKSC 220

QY 181 IEIILAE 187
Db 221 IEIILAE 227

RESULT 11

```

US-10-201-292-28
Sequence 28, Application US/10201292
Publication No. US20030144193A1
GENERAL INFORMATION:
APPLICANT: James B. ROTHMAN
APPLICANT: Theresa L. O'KEEFE
APPLICANT: Engin OZKAYNAK
APPLICANT: Judith J. HEALEY
TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
FILE REFERENCE: 7853-253-999
CURRENT APPLICATION NUMBER: US/10/201,292
CURRENT FILING DATE: 2003-02-14
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FaastSeq for Windows Version 4.0
SEQ ID NO 28
LENGTH: 460
TYPE: PRT
ORGANISM: Homo sapiens
US-10-201-292-28

Query Match 100.0%; Score 970; DB 14; Length 460;
Best Local Similarity 100.0%; Pred. No. 9, 6e-93;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFDFYILDKSGSVLHNMNEIYFVEQLAHKFI SPOLRMSFIVFSTRGTTLMKLTEDRE 60
DB 41 GGFDFYILDKSGSVLHNMNEIYFVEQLAHKFI SPOLRMSFIVFSTRGTTLMKLTEDRE 100
QY 61 QIROGJLELOKVLPGGDTYHHEGFERASBOIYYENRGYRTASVITLTTGELHEDLFFY 120
DB 101 QIROGJLELOKVLPGGDTYHHEGFERASBOIYYENRGYRTASVITLTTGELHEDLFFY 160
QY 121 SEREANSRDLGAIYVCVGVDFNFTQARIADSKDHFVFNDFGFOALOGIHSILKXSC 180
DB 161 SEREANSRDLGAIYVCVGVDFNFTQARIADSKDHFVFNDFGFOALOGIHSILKXSC 220
QY 181 IETILAE 187
DB 221 IETILAE 227

RESULT 12
US-10-201-292-32
Sequence 32, Application US/10201292
Publication No. US20030144193A1
GENERAL INFORMATION:
APPLICANT: James B. ROTHMAN
APPLICANT: Theresa L. O'KEEFE
APPLICANT: Engin OZKAYNAK
APPLICANT: Judith J. HEALEY
TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
FILE REFERENCE: 7853-253-999
CURRENT APPLICATION NUMBER: US/10/201,292
CURRENT FILING DATE: 2003-02-14
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FaastSeq for Windows Version 4.0
SEQ ID NO 32
LENGTH: 479
TYPE: PRT
ORGANISM: Homo sapiens
US-10-201-292-32

Query Match 100.0%; Score 970; DB 14; Length 479;
Best Local Similarity 100.0%; Pred. No. 1e-92;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFDFYILDKSGSVLHNMNEIYFVEQLAHKFI SPOLRMSFIVFSTRGTTLMKLTEDRE 60
DB 41 GGFDFYILDKSGSVLHNMNEIYFVEQLAHKFI SPOLRMSFIVFSTRGTTLMKLTEDRE 100
QY 61 QIROGJLELOKVLPGGDTYHHEGFERASBOIYYENRGYRTASVITLTTGELHEDLFFY 120
DB 101 QIROGJLELOKVLPGGDTYHHEGFERASBOIYYENRGYRTASVITLTTGELHEDLFFY 160

QY 121 SEREANSRDLGAIYVCVGVDFNFTQARIADSKDHFVFNDFGFOALOGIHSILKXSC 180
DB 161 SEREANSRDLGAIYVCVGVDFNFTQARIADSKDHFVFNDFGFOALOGIHSILKXSC 220
QY 181 IETILAE 187
DB 221 IETILAE 227

RESULT 13
US-10-201-292-34
Sequence 34, Application US/10201292
Publication No. US20030144193A1
GENERAL INFORMATION:
APPLICANT: James B. ROTHMAN
APPLICANT: Theresa L. O'KEEFE
APPLICANT: Engin OZKAYNAK
APPLICANT: Judith J. HEALEY
TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
FILE REFERENCE: 7853-253-999
CURRENT APPLICATION NUMBER: US/10/201,292
CURRENT FILING DATE: 2003-02-14
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FaastSeq for Windows Version 4.0
SEQ ID NO 34
LENGTH: 504
TYPE: PRT
ORGANISM: Homo sapiens
US-10-201-292-34

Query Match 100.0%; Score 970; DB 14; Length 504;
Best Local Similarity 100.0%; Pred. No. 1, 1e-92;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFDFYILDKSGSVLHNMNEIYFVEQLAHKFI SPOLRMSFIVFSTRGTTLMKLTEDRE 60
DB 41 GGFDFYILDKSGSVLHNMNEIYFVEQLAHKFI SPOLRMSFIVFSTRGTTLMKLTEDRE 100
QY 61 QIROGJLELOKVLPGGDTYHHEGFERASBOIYYENRGYRTASVITLTTGELHEDLFFY 120
DB 101 QIROGJLELOKVLPGGDTYHHEGFERASBOIYYENRGYRTASVITLTTGELHEDLFFY 160
QY 121 SEREANSRDLGAIYVCVGVDFNFTQARIADSKDHFVFNDFGFOALOGIHSILKXSC 180
DB 161 SEREANSRDLGAIYVCVGVDFNFTQARIADSKDHFVFNDFGFOALOGIHSILKXSC 220
QY 181 IETILAE 187
DB 221 IETILAE 227

RESULT 14
US-10-201-292-36
Sequence 36, Application US/10201292
Publication No. US20030144193A1
GENERAL INFORMATION:
APPLICANT: James B. ROTHMAN
APPLICANT: Theresa L. O'KEEFE
APPLICANT: Engin OZKAYNAK
APPLICANT: Judith J. HEALEY
TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
FILE REFERENCE: 7853-253-999
CURRENT APPLICATION NUMBER: US/10/201,292
CURRENT FILING DATE: 2003-02-14
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FaastSeq for Windows Version 4.0
SEQ ID NO 36
LENGTH: 529
TYPE: PRT
ORGANISM: Homo sapiens
US-10-201-292-36

```
Query Match      100.0%; Score 970; DB 14; Length 529;
Best Local Similarity 100.0%; Pred. No. 1.2e-92;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGFDLVFIIDKSGSVLHHNNEIYYFVEQLAHKFISPOLRMSFIVFSTRGTTLMKLTEDRE 60
DB      41 GGFDLVFIIDKSGSVLHHNNEIYYFVEQLAHKFISPOLRMSFIVFSTRGTTLMKLTEDRE 100

QY      61 QIROGLEELQKVLPGGDTYMHGEPERASEQIYYENRQGYRTASVIALTDGELHEDLFFY 120
DB      101 QIROGLEELQKVLPGGDTYMHGEPERASEQIYYENRQGYRTASVIALTDGELHEDLFFY 160

QY      121 SEREANSRDLGAIYVCVGVDFNETQLARIADSKDHVPVNDGFQALOGIISHILKSC 180
DB      161 SEREANSRDLGAIYVCVGVDFNETQLARIADSKDHVPVNDGFQALOGIISHILKSC 220

QY      181 IEIILAE 187
DB      221 IEIILAE 227

RESULT 15
US-10-038-307-18
; Sequence 18; Application US/10038307
; Publication No US20030134786A1
; GENERAL INFORMATION:
; APPLICANT: James B. ROTTMAN
; APPLICANT: Theresa L. O'KEEFE
; APPLICANT: Engin OZKAYNAK
; APPLICANT: Judith J. HEALEY
; TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
; FILE REFERENCE: 7853-253-999
; CURRENT APPLICATION NUMBER: US/10/038.307
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PasteSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 551
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-038-307-18

Query Match      100.0%; Score 970; DB 14; Length 551;
Best Local Similarity 100.0%; Pred. No. 1.2e-92;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGFDLVFIIDKSGSVLHHNNEIYYFVEQLAHKFISPOLRMSFIVFSTRGTTLMKLTEDRE 60
DB      41 GGFDLVFIIDKSGSVLHHNNEIYYFVEQLAHKFISPOLRMSFIVFSTRGTTLMKLTEDRE 100

QY      61 QIROGLEELQKVLPGGDTYMHGEPERASEQIYYENRQGYRTASVIALTDGELHEDLFFY 120
DB      101 QIROGLEELQKVLPGGDTYMHGEPERASEQIYYENRQGYRTASVIALTDGELHEDLFFY 160

QY      121 SEREANSRDLGAIYVCVGVDFNETQLARIADSKDHVPVNDGFQALOGIISHILKSC 180
DB      161 SEREANSRDLGAIYVCVGVDFNETQLARIADSKDHVPVNDGFQALOGIISHILKSC 220

QY      181 IEIILAE 187
DB      221 IEIILAE 227
```

Search completed: June 13, 2005, 20:36:50
Job time : 59.9593 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 19:47:08 ; Search time 13.4163 Seconds

(without alignments)
1341.095 Million cell updates/sec

Title: US-09-970-076-2_COPY_41_227

Perfect score: 970
Sequence: 1 GGFGLYFLDKSGSVLHMN.....LOGIHSILKSCIEILAE 187Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	137	14.1	724	2 A48569	antigen Em100 - E1
2	132.5	13.7	1153	2 RWHUB	cell surface glyco
3	129	13.3	3124	2 A40020	collagen alpha 1(X
4	128.5	13.2	712	2 A45638	immunodominant mic
5	125.5	12.9	1163	1 RWHUIC	cell surface glyco
6	124.5	12.8	1153	2 S00551	leukocyte surface
7	118.5	12.2	760	1 C2MS	classical-compleme
8	116	12.0	3051	2 S42373	hypothetical prote
9	113	11.6	1747	2 A45974	collagen alpha 1(X
10	113	11.6	1857	2 S31212	collagen alpha 1(X
11	113	11.6	1888	2 S78476	collagen alpha 1(X
12	111.5	11.5	764	1 BBHU	complement factor
13	111.5	11.5	929	2 S51027	type XII collagen
14	110	11.3	2944	2 A54849	collagen alpha 1(V
15	109.5	11.3	1029	1 S21369	collagen alpha 2(V
16	103.5	10.7	567	2 T28797	hypothetical prote
17	103.5	10.7	761	1 BBMS	complement factor
18	100	10.3	460	2 T23087	hypothetical prote
19	99	10.2	698	2 B96958	dnak protein (heat
20	99	10.2	917	2 S09646	collagen alpha 2(V
21	99	10.2	1018	1 CGH2A	collagen alpha 2(V
22	98	10.1	741	2 T46488	hypothetical prote
23	97.5	10.1	496	2 A37979	cartilage matrix p
24	96.5	9.9	493	2 A33809	cartilage matrix p
25	96.5	9.9	574	2 A46283	sporozoite surface
26	95.5	9.8	500	2 S66522	cartilage matrix p
27	95.5	9.8	559	2 S04531	thrombospondin-rel
28	95.5	9.8	1179	2 AS3213	integrin alpha-B-C
29	94	9.7	3176	2 CGH3A	collagen alpha 3(V

30	93.5	9.6	843	2 A40970	undulin 1 - human
31	93	9.6	340	2 E70121	hypothetical prote
32	92.5	9.5	3137	2 A37797	collagen alpha 3(V
33	88.5	9.1	932	2 JC5953	inter-alpha-inhibi
34	88	9.1	1218	2 T30889	valine-tRNA ligase
35	87.5	9.0	341	2 T32949	hypothetical prote
36	87.5	9.0	418	2 AG2350	hypothetical prote
37	87	9.0	710	2 AG2285	hypothetical prote
38	85.5	8.8	340	2 B83262	hypothetical prote
39	85.5	8.8	380	2 H90261	hypothetical prote
40	85	8.8	1021	2 T05901	hypothetical prote
41	84.5	8.7	689	2 F84811	probable retroelem
42	84.5	8.7	800	2 AB1079	endoglucanase homo
43	84.5	8.7	875	1 ITECAP	DNA topoisomerase
44	84.5	8.7	875	2 B91018	DNA gyrase subunit
45	84.5	8.7	875	2 D85862	DNA gyrase, subunit

ALIGNMENTS

RESULT 1

A48569
antigen Em100 - Eimeria maxima

C/Species: Eimeria maxima

C/Date: 01-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

C/Accession: A48569

R:Pasamonres, L.; Hug, D.; Humbelin, M.; Weber, G.
Mol. Biochem. Parasitol. 57, 171-174, 1993

A/Title: Sequence of a major Eimeria maxima antigen homologous to the Eimeria tenella m

A/Reference number: A48569; M01D:93149203; PMID:8426611

A/Accession: A48569

A/Status: preliminary

A/Molecule type: nucleic acid

A/Residues: 1-724 <PMS>

A/Cross-references: UNIPROT:Q04588; GB:M99058; NID:G158890; PTD:G158891

A/Note: sequence extracted from NCBI backbone (NCBI:123776, NCBI:P123777)

F/45-218/Domain: von Willebrand factor type A repeat homology <VWA2>

F/238-296/Domain: thrombospondin type 1 repeat homology <THR1>

F/309-371/Domain: thrombospondin type 1 repeat homology <THR2>

F/372-432/Domain: thrombospondin type 1 repeat homology <THR3>

F/433-493/Domain: thrombospondin type 1 repeat homology <THR4>

F/494-556/Domain: thrombospondin type 1 repeat homology <THR5>

F/560-610/Domain: thrombospondin type 1 repeat homology <THR6>

Query Match 14.1%; Score 137; DB 2; Length 724;

Best Local Similarity 26.6%; Pred. No. 0.0007; Matches 53; Conservative 38; Mismatches 80; Indels 28; Gaps 12;

QY	4	DLVEFLDKSGSV-LHMHNEIYFVEQLAHKF-ISP-QLRMSFTVSTRTLMKLTRE	60
DB	47	DWMLVDESSIGTSNGKRSFISNFGCTMPSPDVRGVLTFGSAVTRMDLS	106
QY	61	QIRQGLEELQKLP-P-GGDTYMEGFERASEQIYENRQGYR-TASVIALTDG	116
DB	107	QNMDLAAAKKLPYAAAGSTYTLGLAKA-EELIFSPKGRDVAPEMLVTMDGA	161
QY	117	LFYFSERE-----ANRSRLGAIYVCVQ-KQFNELQARIA-DKDNV-PF-V	164
DB	162	-----SSRSQTLNAEKLRNRGVIIVLVGVTGVSNAECSRIAGCDTSYV	217
QY	165	FOALOGIHSILKSCIEI 183	
DB	218	WGVSSQINIGIRKACDL 236	

RESULT 2

RWHUB
cell surface glycoprotein CD11b precursor (validated) - human

N/Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein Mac

eukocyte integrin alpha chain; neutrophil adherence receptor alpha chain

C/Species: Homo sapiens (man)

C/Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004

C/Accession: A31108; A28915; A41600; A30892; A32218; A46526; A26091; I52567
 R/Corb, A.L.; Kishimoto, T.K.; Miller, U.J.; Springer, T.A.
 J. Biol. Chem. 263, 12403-12411, 1988
 A/Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD
 B.
 A/Reference number: A31108; MUID:88315033; PMID:2457584
 A/Accession: A31108
 A/Molecule type: mRNA
 A/Residues: 1-1153 <COR>
 A/Cross-references: UNIPROT:P11215; GB:J03925; NID:g187284; PIDN:AAA59544.1; PID:g307148
 A/Note: part of this sequence was confirmed by protein sequencing
 R/Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.
 J. Cell Biol. 106, 2153-2158, 1988
 A/Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor M
 A/Reference number: A28915; MUID:88257215; PMID:2454931
 A/Accession: A28915
 A/Molecule type: mRNA
 A/Residues: 1-499,501-965,'P',967-1153 <ARN>
 A/Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:g186935; PIDN:AAA594
 A/Note: the authors translated the codon TAC for residue 1129 as Thr
 R/Shelley, C.S.; Arnaout, M.A.
 Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991
 A/Title: The promoter of the CD1b gene directs myeloid-specific and developmentally reg
 A/Reference number: A41600; MUID:92073318; PMID:1683702
 A/Accession: A41600
 A/Molecule type: DNA
 A/Residues: 1-9 <SHE>
 A/Cross-references: GB:M76724; NID:g180018; PIDN:AAA58410.1; PID:g553215
 R/Arnaout, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.
 Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988
 A/Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhe
 A/Reference number: A94193; MUID:88190151; PMID:2833753
 A/Accession: A30892
 A/Molecule type: mRNA
 A/Residues: 917-1042 <AR2>
 A/Cross-references: GB:M18044
 R/Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.
 Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989
 A/Title: CDNA sequence for the alpha subunit of the human neutrophil adherence receptor
 A/Reference number: A32218; MUID:89098893; PMID:2563162
 A/Accession: A32218
 A/Molecule type: mRNA
 A/Residues: 9-1153 <HIC>
 A/Cross-references: GB:U04145; NID:g189068; PIDN:AAA5903.1; PID:g386975
 A/Note: part of this sequence was confirmed by protein sequencing
 R/Pleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.
 J. Immunol. 150, 480-490, 1993
 A/Title: Structural analysis of the CD1b gene and phylogenetic analysis of the alpha-1
 n during evolution.
 A/Reference number: A46526; MUID:93123748; PMID:8419480
 A/Accession: A46526
 A/Status: not compared with conceptual translation
 A/Molecule type: DNA
 A/Residues: 1-499,501-1153 <PLE>
 A/Cross-references: GB:S52227; NID:g263047; PIDN:AA24821.1; PID:g263049
 A/Note: the last three bases of Intron 13, CAG, are included in some but not all mature
 A/Note: sequence extracted from NCBI backbone (NCBI:P121963)
 R/Pierce, M.W.; Remold-O'Donnell, E.; Todd III, R.F.; Arnaout, M.A.
 Biochim. Biophys. Acta 874, 368-371, 1986
 A/Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across sp
 A/Reference number: A90664; MUID:87076671; PMID:3539202
 A/Accession: A90664
 A/Molecule type: protein
 A/Residues: 17-31 <PIE>
 A/Experimental source: granulocytes
 R/Pahl, H.L.; Rosemarin, A.G.; Tenen, D.G.
 Blood 79, 865-870, 1992
 A/Title: Characterization of the myeloid-specific CD1b promoter.
 A/Reference number: I52567; MUID:92144986; PMID:1346576
 A/Accession: I52567
 A/Status: translated from GB/EMBL/DBJ
 A/Molecule type: DNA

A/Residues: 1-9 <RES>
 A/Cross-references: GB:M4477; NID:g180184; PIDN:AAA51960.1; PID:g553219
 C/Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1 c
 A/Genetics:
 A/Gene: GDB:ITGAM, CR3A
 A/Cross-references: GDB:120599; OMIM:120980
 A/Map position: 16p11.2-16p11.2
 A/Note: promoter contains a GATA motif and two Sp1 consensus binding sites
 C/Superfamily: cell surface glycoprotein CD1b; von Willebrand factor type A repeat homo
 C/Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; magr
 F:1-16/Domain: signal sequence #status predicted <SIG>
 F:17-1153/Product: cell surface glycoprotein CD1b #status experimental <MNT>
 F:17-1108/Domain: extracellular #status predicted <EXT>
 F:18-318/Domain: von Willebrand factor type A repeat homology <VWA2>
 F:465-473/Region: calcium/magnesium binding #status predicted
 F:530-538/Region: calcium/magnesium binding #status predicted
 F:593-601/Region: calcium/magnesium binding #status predicted
 F:1109-1134/Domain: transmembrane #status predicted <TM>
 F:1135-1153/Domain: intracellular #status predicted <INT>
 F:86,240,391,469,693,697,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Binding
 Query Match 13.7%; Score 132.5; DB 1; Length 1153;
 Best Local Similarity 26.3%; Pred. No. 0.003;
 Matches 52; Conservative 41; Mismatches 66; Indels 39; Gaps 11;
 QY 4 DLYFIIDKSGSVV-HHNNELYYVEQLAKFISPOLMSEIVSTGRTIMKLTED----58
 DB 150 DIAFLIDSGSISIPHDFFRKKEFVST---VNEQLKSKSTLPS-----IMQYSEPRRIH 199
 QY 59 -----RQIIGGLEELQKVPFGDPTVWHEGFERSAQIYENRGVYTA-SVIRALT 109
 DB 200 FTFKKQNNPNPNSLVKPIQTQL--GRTHATGIRKVVRELPIITGARGAKKILVIT 257
 QY 110 DGEIHEDLFFYSE--RANKSRDLGAIYVCVQKDFNETQLAR----IADS--KDHVPF 160
 DB 258 DGEKFDGPGYEDVIDEADRE---GVIRYVIGVGDAPFRSKRQELNTATSKPPRDHVFQ 314
 QY 161 VNDGFQALQGIHTSIILK 178
 DB 315 VNN-FEALKTIONQLREK 331
 RESULT 3
 A40020
 Collagen alpha 1(XII) chain precursor - chicken
 N/Alternate names: fibrochimerin
 C/Species: Gallus gallus (chicken)
 C/Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 09-Jul-2004
 C/Accession: A40020; A34485; B34485; A28037; S23814; S22554; S28811
 R/Yamagata, M.; Yamada, K.M.; Yamada, S.S.; Shimomura, T.; Tanaka, H.; Nishida, Y.; Obari
 J. Cell Biol. 115, 209-221, 1991
 A/Title: The complete primary structure of type XII collagen shows a chimeric molecule w
 nou region of type IX collagen, and short collagenous domains with an Arg-Gly-Asp site.
 A/Reference number: A40020; MUID:92011862; PMID:1918137
 A/Accession: A40020
 A/Molecule type: mRNA
 A/Residues: 1-3124 <YAM>
 A/Cross-references: UNIPROT:P13944; GB:D00824; NID:g222810; PIDN:BA400701.1; PID:g222811
 A/Note: in the authors' translation residues 1216-1219 are shown after residue 1235 and,
 R/Gordon, M.K.; Gerecke, D.R.; Dubble, B.; van der Rest, M.; Olsen, B.R.
 J. Biol. Chem. 264, 19772-19778, 1989
 A/Title: Type XII collagen. A large multidomain molecule with partial homology to type I
 A/Reference number: A34485; MUID:9062079; PMID:2584192
 A/Accession: A34485
 A/Molecule type: mRNA
 A/Residues: 2456-2758 'A', 2760-2802, 'F', 2804-2976, 'F', 2978-3124 <GOR>
 A/Cross-references: EMBL:U05137; NID:g211284; PIDN:AAA46835.1; PID:g211285
 A/Accession: B34485
 A/Molecule type: protein
 A/Residues: 2772-2792/2846-2873 <GOR2>
 R/Gordon, M.K.; Gerecke, D.R.; Olsen, B.R.
 Proc. Natl. Acad. Sci. U.S.A. 84, 6040-6044, 1987
 A/Title: Type XII collagen: distinct extracellular matrix component discovered by cDNA c]

A/Reference number: A36584
 A/Contents: erratum
 A/Accession: A36584
 A/Molecule type: DNA
 A/Residues: 1-1163 <COR>
 A/Cross-references: UNIPROT:P20702
 A/Note: this revision to the sequence from reference A35543 includes the carboxyl end
 R/Corbl. A.L.; Garcia-Aguilar, J.; Springer, T.A.
 J. Biol. Chem. 265, 2782-2788, 1990
 A/Title: Genomic structure of an integrin alpha subunit, the leukocyte p150,95 molecule.
 A/Reference number: A35543; MUID:90153906; PMID:2303426
 A/Accession: A35543
 A/Molecule type: DNA
 A/Residues: 1-834 <CO2>
 A/Note: this sequence has been revised in reference A36584
 R/Corbl. A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.
 EMBO J. 6, 4033-4028, 1987
 A/Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocyte
 A/Reference number: 500664; MUID:88166645; PMID:3327687
 A/Accession: 500664
 A/Molecule type: mRNA
 A/Residues: 1-755; 'L', 757-1163 <CO3>
 A/Cross-references: GB:M81695; EMBL:Y00093; NID:G487829; PIDN:AAA59180.1; PID:G487830
 A/Note: part of this sequence was confirmed by protein sequencing
 C/Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on my
 C/Genetics: GDB:ITGAX; CD11C
 A/Gene: GDB:ITGAX; CD11C
 A/Cross-references: GDB:119758; OMIM:151510
 A/Map position: 16p11.2-16p11.2
 C/Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
 C/Keywords: calcium; cell adhesion; glycoprotein; heterodimer; magnesium; tandem repeat;
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-1163/Product: cell surface glycoprotein CD11c #status predicted <MAT>
 F:149-319/Domain: extracellular #status predicted <EXT>
 F:1108-1133/Domain: von Willebrand factor type A repeat homology <VMA4>
 F:1134-1163/Domain: transmembrane #status predicted <TM>
 F:61,89,392,657,735,899,939,1050/Binding site: carbohydrate (Asn) (covalent) #status pre
 Query Match 12.9%; Score 125.5; DB 1; Length 1163;
 Best Local Similarity 24.7%; Pred. No. 0.012;
 Matches 48; Conservative 38; Mismatches 77; Indels 31; Gaps 9;
 QY 4 DLVFLDKSGSV-LHMNEIYFVEQLAHKFTISQLMSPFVSTRTGTTMLKLTEDP-- 60
 DB 151 DIVFLIDSGSISRNPAWNNFRAVISOFPORSTQFSIMQSNKQTFTEPEEFRTS 210
 QY 61 ---QIRQGLELQKVLPGSDTYMHGEFERSAQIYYENRQGYRTAS-VITATDGEIHED 116
 DB 211 NPISLASVHQLQ-----GFTYTAIAQNVVHRLFHASYGARRDATKILVITDGKKEGD 265
 QY 117 LFVYSERENSRDLGAIYVCVG-----KDFNETQLARIAD--SKDHVPVNDG 164
 DB 266 SLDYKD-VIMADAAGITRAIVGIGLAFQNRNMSKEIND-----IASKPSQEHIFKED- 318
 QY 165 FQALQGIHHSILKK 178
 DB 319 FDLAKDIONQLKEX 332
 RESULT 6
 S00551
 Leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse
 N/Alternate names: complement-3 receptor alpha chain
 C/Species: Mus musculus (house mouse)
 C/Date: 30-Sep-1989 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
 C/Accession: S00551; I55078
 R/Pytela, R.
 EMBO J. 7, 1371-1378, 1988
 A/Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the
 A/Reference number: S00551; MUID:88312584; PMID:3044779
 A/Accession: S00551
 A/Molecule type: DNA

A/Residues: 1-1153 <PYT>
 A/Cross-references: UNIPROT:P05555; EMBL:X07640; NID:952982; PIDN:CAA30479.1; PID:g552983
 A/Note: the authors translated the codon CAC for residue 569 as Gln
 R/Sastre, L.; Roman, J.M.; Teplov, D.B.; Dreyer, W.J.; Gae, C.E.; Larson, R.S.; Roberts,
 Proc. Natl. Acad. Sci. U.S.A. 83, 5644-5648, 1986
 A/Title: A partial genomic DNA clone for the alpha subunit of the mouse complement recept
 A/Reference number: I59078; MUID:86287312; PMID:2942940
 A/Accession: I59078
 A/Status: preliminary; translated from GB/EMBL/DD8J
 A/Molecule type: DNA
 A/Residues: 11-44 <RSS>
 A/Cross-references: GB:M14293; NID:g198993; PIDN:AAA39484.1; PID:g554193
 C/Genetics:
 A/Gene: Mac-1
 C/Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homo
 C/Keywords: cell adhesion; glycoprotein; transmembrane protein
 F:1-16/Domain: signal sequence #status predicted <SIG>
 F:17-1153/Product: leukocyte surface glycoprotein Mac-1 alpha chain #status experimental
 F:148-118/Domain: von Willebrand factor type A repeat homology <VMA2>
 F:1106-1123/Domain: transmembrane #status predicted <TM>
 Query Match 12.8%; Score 124.5; DB 2; Length 1153;
 Best Local Similarity 26.8%; Pred. No. 0.014;
 Matches 53; Conservative 37; Mismatches 69; Indels 39; Gaps 12;
 QY 4 DLVFLDKSGSV-LHMNEIYFVEQLAHKFTIS--PQLRMSPFVSTRTGTTMLKLTED-- 58
 DB 150 DIVFLIDSGSSI-----NNIDF--QKWKERFVSTVMEQFKSKTLFS-----LMQYSDPFR 197
 QY 59 ---REQIRGLELQKVLPG---GGDTVMHGEFERSAQIYYE-NRQGYRTASVITAT 109
 DB 198 IHTFTNDPKRNPSPRSHSVSYKQNGRTKTSAIRKYVRELFKINGCARNAKKILVIT 257
 QY 110 DGEIHEDLFPEYS--BEANRSRLGAIYVCVGKD--FNETQLARIAD-----SKDHVP 160
 DB 258 DGEKFGDPLDYKQVTEADRA---GIRYIVIGVGNAPNKRQSRRLDTIASKPAGENVFQ 314
 QY 161 VNDGFQALQGIHHSILKK 178
 DB 315 V-DNFEALNTIQNLQLEK 331
 RESULT 7
 C2N8
 Classical-complement-pathway C3/C5 convertase (EC 3.4.21.43) C2 component precursor - mo
 N/Alternate names: C3 convertase; C5 convertase; complement C2
 C/Species: Mus musculus (house mouse)
 C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
 C/Accession: A38876; B36593; I54429
 R/Ishikawa, N.; Nonaka, M.; Wetzel, R.A.; Colten, H.R.
 submitted to Genbank, January 1991
 A/Reference number: A38875
 A/Accession: A38876
 A/Molecule type: DNA
 A/Residues: 1-760 <IS2>
 A/Cross-references: UNIPROT:P21180; GB:M57891; GB:J05661; NID:g192436; PIDN:AAA63294.1; I
 R/Ishikawa, N.; Nonaka, M.; Wetzel, R.A.; Colten, H.R.
 J. Biol. Chem. 265, 19040-19046, 1990
 A/Title: Murine complement C2 and factor B genomic and cDNA cloning reveals different me
 A/Reference number: A36593; MUID:91035430; PMID:2229060
 A/Accession: B36593
 A/Molecule type: mRNA
 A/Residues: 1-760 <ISH>
 A/Cross-references: EMBL:M57891; NID:g192436; PIDN:AAA63294.1; PID:g192437
 R/Palus, A.; Wakeland, E.K.; McConnell, T.J.; Gitlin, J.; Whitehead, A.S.; Colten, H.R.
 Immunogenetics 25, 290-298, 1987
 A/Title: DNA polymorphism of MHC III genes in inbred and wild mouse strains.
 A/Reference number: I54429; MUID:87192938; PMID:2883115
 A/Accession: I54429
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 660-677; 'R', 679, 681-723; 'G', 725 <RSS>
 A/Cross-references: GB:M16271; NID:g199289; PIDN:AAA9562.1; PID:g199290

A/Reference number: A44622; MUID:83039428; PMID:6957884
A/Accession: A44622
A/Molecule type: mRNA
A/Residues: 467-546/550-595/752-764 <MO>
A/Cross-references: GB:U00185; GB:U00186
A/Note: The authors translated the codon TAC at 519 as Thr: the nucleic acid translation
R/Mole: J.E.; Anderson, J.K.; Davidson, E.A.; Woods, D.E.
J. Biol. Chem. 259, 3407-3412, 1984
A/Title: Complete primary structure for the zymogen of human complement factor B.
A/Reference number: A20751; MUID:84161997; PMID:6546754
A/Accession: A00934
A/Molecule type: Protein; mRNA
A/Residues: 26-764 <MO>
A/Cross-references: GB:K01566
A/Note: nucleic acid translation differs from the sequence shown in having 300-Leu, 328-
A/Note: glycosylation sites were determined
R/Chetani, D.L.; Gagnon, J.
Biochem. J. 209, 61-70, 1983
A/Title: Amino acid sequence of the Bb fragment from complement factor B. Sequence of th
A/Reference number: A19188; MUID:83204002; PMID:6342610
A/Contents: the final paper in a series documenting the sequence, glycosylation site, at
A/Accession: A19188
A/Molecule type: Protein
A/Residues: 260-286/298-764 <CHR>
R/Campbell, R.D.; Porter, R.R.
Proc. Natl. Acad. Sci. U.S.A. 80, 4464-4468, 1983
A/Title: Molecular cloning and characterization of the gene coding for human complement
A/Reference number: A19947; MUID:83273641; PMID:6308626
A/Accession: A19947
A/Molecule type: DNA
A/Residues: 346-764 <CAM>
A/Cross-references: GB:U00125
A/Accession: B19947
A/Molecule type: mRNA
A/Residues: 339-509 <CA1>
A/Cross-references: GB:U00126; NID:9187723; PIDN:AAA36226.1; PID:9553536
R/Wu, L.; Morley, B.J.; Campbell, R.D.
Cell 48, 331-342, 1987
A/Title: Cell-specific expression of the human complement protein factor B gene: evidence
A/Reference number: A25971; MUID:87102880; PMID:3643061
A/Accession: B25971
A/Molecule type: DNA
A/Residues: 1-99 <MDU>
A/Cross-references: GB:M15082; NID:9187699; PIDN:AAA59625.1; PID:9553534
R/Niemann, M.A.; Bhown, A.S.; Miller, E.J.
Biochem. J. 274, 473-480, 1991
A/Title: The principal site of glycation of human complement factor B.
A/Reference number: S14339; MUID:91174758; PMID:2006911
A/Accession: S14339
A/Molecule type: Protein
A/Residues: 270-329 <NIE>
A/Note: binding site for carbohydrate to lysine under artificial conditions
R/Morley, B.J.; Campbell, R.D.
EMBO J. 3, 153-157, 1984
A/Title: Internal homologues of the Ba fragment from human complement component factor B
A/Reference number: A44628; MUID:84158524; PMID:6323161
A/Accession: A44628
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 16-225/227-259 <MOR>
R/Schwede, W.; Lutting, B.; Sokolowski, T.; Estaller, C.; Weiss, E.H.; Meyer zum Busche
Immunobiology 188, 221-232, 1993
A/Title: Human complement factor B: functional properties of a recombinant zymogen of th
A/Reference number: I54409; MUID:94041399; PMID:8225386
A/Accession: I54409
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-764 <RES>
A/Cross-references: GB:667310; NID:9452937; PIDN:AA13989.1; PID:94261689
R/Horuchi, T.; Kim, S.; Matsunoto, M.; Watanabe, I.; Fujita, S.; Volanakis, J.E.
Mol. Immunol. 30, 1567-1592, 1993
A/Title: Human complement factor B: CDNA cloning, nucleotide sequencing, phenotypic con

A/Reference number: I57824; MUID:94067177; PMID:8247029
A/Accession: I57824
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-31/0/33-764 <RB2>
A/Cross-references: GB:U15702; NID:9291921; PIDN:AAA16820.1; PID:9291922
C/Comment: 292-Cys has a free sulfhydryl.
C/Genetics:
A/Gene: GDB:BF
A/Cross-references: GDB:119726; OMIM:138470
A/Map position: 6p21.3-6p21.3
A/Introns: 21/3; 99/3; 346/1; 390/1; 424/1; 470/1; 502/3; 542/1; 593/2; 619/1; 652/3; 69
A/Note: the list of introns may be incomplete
A/Note: gene is located in the major histocompatibility complex, class III region
C/Complex: complement factor B initially forms an inactive complex with complement facto
ment factor C3b forming active C3/C5 convertase; Ba is released
C/Function:
A/Description: Bb is a serine proteinase; C3/C5 convertase cleaves complement C3 alpha c
a1
A/Pathway: complement alternate pathway
C/Superfamily: complement B/C2; complement factor H repeat homology; trypsin homology; v
C/Keywords: acute phase; complement alternate pathway; duplication; glycoprotein; hydrol
P:1-35/Domain: signal sequence #status predicted <SIG>
P:26-764/Product: complement factor B #status experimental <MAT>
P:26-259/Product: complement factor Ba fragment #status experimental <BAF>
P:37-98/Domain: complement factor H repeat homology <FH1>
P:103-158/Domain: complement factor H repeat homology <FH2>
P:165-218/Domain: complement factor H repeat homology <FH3>
P:260-764/Product: C3/C5 convertase Bb fragment #status experimental <BBF>
P:268-458/Domain: von Willebrand factor type A repeat homology <VFA>
P:482-752/Domain: trypsin homology #status atypical <TRY>
P:37-76/62-98, 103-145, 131-158, 165-205, 191-218, 478-596, 511-527, 599-615, 656-682, 695-725/DI
F:122,142,285,378/Binding site: carbohydrate (Aan) (covalent) #status experimental
F:259-260/Cleavage site: Arg-Lys (complement factor D) #status experimental
F:526,576,699/Active site: His, Asp, Ser #status experimental
Query Match 11.5%; Score 111.5; DB 1; Length 764;
Best local similarity 22.8%; Pred. No. 0.11;
Matches 49; Conservative 44; Mismatches 77; Indels 45; Gaps 10;
Qy 1 GGFPLRYILDKSGV-----LHMNELYFPEQLAHFISQLMSRIVSTRTGLTKL 55
Db 267 GSMNIYLVLDGSSISIGASNFTGAKKCLVNLIEKVASGVXPR--RYGLVYATYATPKIWKV 324
Qy 56 TE-----DREQIRQLLEL-----OKVLPGDPYVMEGPERSEQIYENR-----OGY- 99
Db 325 SEADSSNADWYTKOLMEINTEBHKLSGCTNT-----KKALQAVYSMMSPDDVPPGWN 378
Qy 100 RTASVITATLDG-----ELHEDLPFYSRBRANRSRDLGAIYCVG--VXDFN 144
Db 379 RTRRVITLMTDGLHNMGGDPITVIDELRDLLYICKKKNRBDLDVYFVGVLVQVN 438
Qy 145 ETQLARIADSKDYFPVNDGFQALGIHSLTKS 179
Db 439 INMLASKDNEQHFVKYD--MENLEDVYQWIDS 472
RESULT 13
151027
type XII collagen alpha-1 chain - eastern newt (fragment)
C/Species: Notothenchomus viridescens, Triturus viridescens (eastern newt)
C/Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C/Accession: I51027
R/Mel, Y.; Yang, E.V.; Klatt, K.P.; Tassava, R.A.
Dev. Biol. 168, 503-513, 1995
A/Title: Monoclonal antibody WT2 identifies the urodele alpha 1 chain of type XII collagen
A/Reference number: I51027; MUID:95246925; PMID:7729585
A/Accession: I51027
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-929 <MEI>
A/Cross-references: UNIPROT:Q91145; EMBL:U19494; NID:9633647; PIDN:AAA80217.1; PID:963264
P:155-236/Domain: fibronectin type III repeat homology <3FR>

F:631-795/Domains: von Willebrand factor type A repeat homology <VWA>

```
Query Match      11.5%   Score 111.5; DB 2; Length 928;  
Beet Local Similarity 23.2%; Pred.No.0.15;  
Matches 43; Conservative 37; Mismatches 90; Indels 15; Gaps 6;  
  
QY 4 DLYFLDSSGSAV-LHMHNIYYFVEQLAHKISIPOLAFSFIYFSTGRGTLLAKLTDEBQI 62  
DB :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
633 DIVLVSWSIGRNPFIVANFISRVIVEVDIGSDRQIVASQSGDPRTMQLNTHKT 692  
QY :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
63 RQGLEELQKLP--GGDTVMHEGFERASEQIYENRGY---RTASYIALTLDELIED 116  
DB :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
693 KKSIMDAVANPYKKGNGNTTGSALKFTLENNP---RPVGMRERARKAIALLTGKSCDD 749  
QY :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
117 LFFSERANSRDGLAVICYGVDPNETQARIADSKHVPFN-DGPALGIIHSI 175  
DB :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
750 IVASRKRYA----DGIELIVAGINADENELKEIASPDDELMTNVADPSLTITIVNDL 805  
QY :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
Db 806 TENVC 810
```

RESULT 14

A54849
collagen alpha 1(VII) chain precursor - human
N.Alternate names: procollagen alpha 1(VII) chain
C.Species: Homo sapiens (man)
C.Date: 04-Nov-1994 #sequence revision 04-Nov-1994 #text change 09-Jul-2004
C.Accession: A54849; PH0844; S16316; I56328; A30296; I64686
R.Christiano, A.M.; Greenspan, D.S.; Lee, S.; Uitto, J.
J. Biol. Chem. 269, 20256-20262, 1994
A.Title: Cloning of human type VII collagen. Complete primary sequence of the alpha 1(VII)
A.Reference number: A54849; MUID:94327588; PMID:8051117
A.Accession: A54849
A.Status: not compared with conceptual translation
A.Molecule type: mRNA
A.Residues: 1-2944 <CHR>
A.Cross-references: UNIPROT:O02388; GB:L02870; NID:9987124; PIDN:AAA75438.1; PID:G987123E
R.Tanaka, T.; Takahashi, K.; Furukawa, F.; Imamura, S.
Biochem. Biophys. Res. Commun. 183, 958-963, 1992
A.Title: Molecular cloning and characterization of type VII collagen cDNA.
A.Reference number: PH0844; MUID:92231902; PMID:1567409
A.Accession: PH0844
A.Molecule type: mRNA
A.Residues: 'ERR', 340-475, 'PALSYASHSLTCWRATRMHPGNRSGSHWTRACPCPNRPASHRAARAG', 524-528, 'C',
A.Cross-references: DDBJ:D11152, DDBJ:D13694; NID:9453698; PIDN:BAA02853.1; PID:9453699
A.Experimental source: keratinocyte
A.Note: the authors translated the codon ACC for residues 394 and 397 as Tyr
R.Patente, M.G.; Chung, L.C.; Rymnenen, J.; Woodley, D.T.; Wynn, K.C.; Bauer, E.A.; Mat
Proc. Natl. Acad. Sci. U.S.A. 88, 6931-6935, 1991
A.Title: Human type VII collagen: cDNA cloning and chromosomal mapping of the gene.
A.Reference number: S16316; MUID:91334380; PMID:1871109
A.Accession: S16316
A.Molecule type: mRNA
A.Residues: 815-892, 'E', 894-1439 <PAR>
A.Cross-references: GB:M6518; GB:S49017; NID:g180914; PIDN:AAA96439.1; PID:g180915
A.Experimental source: Keratinocyte
R.Gammon, W.R.; Abernethy, M.L.; Padilla, K.M.; Pilsavanh, P.S.; Cook, M.E.; Wright, J.
Invest. Dermatol. 99, 691-696, 1992
A>Title: Noncollagenous (NC1) domain of collagen VII resembles multidomain adhesion prob
A.Reference number: I56328; MUID:93107742; PMID:1469284
A.Accession: I56328
A>Status: translated from GB/EMBL/DDBJ
A.Molecule type: mRNA
A.Residues: 'ERR', 372-517, 'DV', 520-540, 'W', 542-1255 <RES>
A.Cross-references: GB:S51236; NID:9262308; PIDN:AAB24637.1; PID:9262309
R.Seltzer, J.L.; Eisen, A.Z.; Bauer, E.A.; Morris, N.P.; Glanville, R.W.; Bugesen, R.E.
J. Biol. Chem. 264, 3822-3826, 1989
A>Title: Cleavage of type VII collagen by interstitial collagenase and type IV collagenena
A.Reference number: A30296; MUID:89139437; PMID:2537292
A.Accession: A30296
A.Molecule type: Protein

[illegible]

Db 210 LVSRRVC 216

RESULT 15

821369

collagen alpha 2(VI) chain precursor - mouse

C/Species: Mus musculus (house mouse)

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C/Accession: S21369; S28808; S13745; S18863

R/Author: X.Y.Z.; Bardon, S.; Amri, E.Z.; Grimaldi, P.; Althaud, G.; Dani, C.

submitted to the EMBL Data Library, April 1992

A/Reference number: S21369

A/Accession: S21369

A/Molecule type: mRNA

A/Residues: 1-1029 <IBR>

A/Cross-references: UNIPROT:Q02788; EMBL:X65582; NID:G49808; PIDN:CAA46541.1; PID:G49808

R/Author: A.; Bertrand, B.; Bardon, S.; Amri, E.Z.; Grimaldi, P.; Althaud, G.; Dani, C.

Biochem. J. 289, 141-147, 1993

A/Title: Cloning of alpha2 chain of type VI collagen and expression during mouse develop

A/Reference number: S28808; PMID:8380980

A/Accession: S28808

A/Molecule type: mRNA

A/Residues: 286-1029 <IB2>

A/Cross-references: EMBL:X62332; NID:G49906; PIDN:CAA44206.1; PID:G49907

R/Author: C.D.; Jimenez, S.A.

Matrix 11, 1-9, 1991

A/Title: Structure of cDNAs encoding the triple-helical domain of murine alpha-2(VI) col

rate oligonucleotides for generation of novel cDNA clones.

A/Reference number: S13745; PMID:1709252

A/Accession: S13745

A/Molecule type: mRNA

A/Residues: 266-267, 'S', 269-294, 'L', 296-600 <CON>

A/Cross-references: GB:U06343; NID:G192671; PIDN:AAA37441.1; PID:G192672

A/Note: the sequence from Fig. 3 is inconsistent with that from Fig. 4 in having 306-116

C/Keywords: alternative splicing; cartilage; cell binding; connective tissue; disulfide

F/1-28/Domain: signal sequence #status predicted <SIG>

F/29-1029/Product: collagen alpha 2(VI) chain #status predicted <MAT>

F/29-265/Domain: globular #status predicted <NC2>

F/54-226/Domain: von Willebrand factor type A repeat homology <VWA1>

F/266-600/Domain: collagenous #status predicted <COL>

F/376-378/Region: cell attachment (R-G-D) motif

F/436-438/Region: cell attachment (R-G-D) motif

F/499-501/Region: cell attachment (R-G-D) motif

F/508-510/Region: cell attachment (R-G-D) motif

F/549-551/Region: cell attachment (R-G-D) motif

F/601-1029/Domain: globular #status predicted <NC1>

F/623-799/Domain: von Willebrand factor type A repeat homology <VWA2>

F/841-1012/Domain: von Willebrand factor type A repeat homology <VWA3>

F/150,337,640,907,963/Binding site: carbohydrate (Aen) (covalent) #status predicted

Query Match 11.3%; Score 109.5; DB 1; Length 1029;

Best Local Similarity 23.5%; Pred. No. 0.24;

Matches 50; Conservative 41; Mismatches 85; Indels 37; Gaps 8;

QY 4 DLYFIDKSGSV-----LHMMN-YYFVEQLAKFTISPOLRMSF-----IVFSTRG 49
 DB 56 NVYFVLDTSSEVAMQSTDSULYHMOQFVQFISQLONEFYLDVALSMRYGGLHFSQV 115
 QY 50 TLMKLTEDREQRQGLLELOKVLPGDLYMHGFERASBOIYYENRQGYRTASVIALT 109
 DB 116 EVFSPGSDASFTKSLQGISFRFG--TFTDCAIANMTQIQRHVGKGVNFAVVI--T 171
 QY 110 DGEIHEDLFPYSEKRNKRDLAGIYCVGV-KDFFETOLARIADSQDHVFPVN----- 162
 DB 172 DGHVTSPPCGGIKQARARBEGRILPALAPNRNINEQGLRIANSPEHLYRNRYATMRP 231
 QY 163 DGFOLQGIHSHIL-----KSCIEI 183
 DB 232 DSTBIODDTINRIKMKHAYGECYKVSCEI 264

Search completed: June 13, 2005, 20:06:37
 Job time : 15.4163 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 19:37:47 ; Search time 62.775 Seconds

(without alignments)
1525.429 Million cell updates/sec

Title: US-09-970-076-2_COPY_41_227

Perfect score: 970
Sequence: 1 GGFGLYFLDKSGSVLHMN.....LOGIHSILKSCIEILAAE 187

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*
1: uniProt_sprot.*
2: uniProt_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	970	100.0	564	1	Q9H6X2 homo sapien
2	961	99.1	562	1	Q9CZ52 mus musculu
3	550	56.7	489	1	P58335 homo sapien
4	548	56.5	487	2	Q6DFX2
5	388	40.0	641	2	Q8BVW2
6	146	15.1	1332	2	Q9BPQ8
7	141.5	14.6	1161	1	ITAD RAT
8	141	14.5	441	2	Q8T6U5
9	141	14.5	441	2	Q8T6U3
10	139	14.3	3119	1	CAIC MOUSE
11	138	14.2	453	2	Q8T5C2
12	137	14.1	724	2	Q04588
13	132.5	13.7	1152	1	ITAM HUMAN
14	131	13.5	3063	1	CAIC HUMAN
15	129.5	13.4	1182	1	Q8C6K9
16	129	13.3	3124	2	CAIC CHICK
17	128.5	13.2	712	2	Q43981
18	127.5	13.1	765	2	Q9U8U9
19	126.5	13.0	517	2	Q43853
20	125.5	12.9	1163	1	ITAM HUMAN
21	124.5	12.8	1153	1	ITAM MOUSE
22	124.5	12.8	1162	1	ITAD HUMAN
23	123	12.7	1626	2	Q8NFU1
24	121.5	12.5	760	2	Q70350
25	121.5	12.5	920	2	Q28984
26	118.5	12.2	760	1	Q02 MOUSE
27	117	12.1	440	2	Q8CBT2
28	117	12.1	846	2	Q8C720
29	117	12.1	1823	2	Q7PRP5
30	117	12.1	3567	2	Q9ES77
31	116.5	12.0	919	2	Q75R52

32	116	12.0	637	2	Q8IVX1	Q8IVX1 homo sapien
33	116	12.0	1284	2	Q6P159	Q6P159 homo sapien
34	116	12.0	1329	1	KE10 HUMAN	Q9P218 homo sapien
35	116	12.0	2944	2	Q63870	Q63870 mus musculu
36	116	12.0	3183	2	Q65ZC2	Q65ZC2 caenorhabdi
37	116	12.0	3767	1	MUA3 CAEEL	P34576 caenorhabdi
38	114.5	11.8	2104	2	Q21281	Q21281 caenorhabdi
39	114.5	11.8	2104	2	Q644N4	Q644N4 caenorhabdi
40	113.5	11.7	790	2	Q6DCQ6	Q6DCQ6 xenopus lae
41	113	11.6	599	2	Q8MVQ1	Q8MVQ1 bolentia vi
42	113	11.6	1868	1	CAIE CHICK	P32018 gallus galli
43	112.5	11.6	1337	2	Q8C972	Q8C972 mus musculu
44	112.5	11.6	1034	2	Q8K229	Q8K229 mus musculu
45	112	11.5	588	2	Q6ZED8	Q6ZED8 synechocyst

ALIGNMENTS

RESULT 1	ATRI HUMAN	STANDARD:	PRT:	564 AA.
ID	Q9H6X2	Q9NPV3:		
AC	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	05-JUN-2004 (Rel. 44, Last annotation update)			
DE	Anthrax toxin receptor 1 precursor (Tumor endothelial marker 8).			
GN	Name=ANTXR1; Synonyms=ATR, TEM8;			
OS	Homo sapiens (Human);			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RX	MEDLINE=20407466; PubMed=10947988; DOI=10.1126/science.289.5482.1197;			
RA	St Croix B., Rago C., Velculescu V., Traverso G., Romans K.E.,			
RA	Montgomery E., Lai A., Riggs G.J., Lengauer C., Vogelstein B.,			
RA	Kinzler K.W.;			
RT	"Gene expressed in human tumor endothelium";			
RL	Science 289:1197-1202(2000).			
RN	[2]			
RP	SEQUENCE FROM N.A. (ISOFORM 2), AND INTERACTION WITH ANTHRAX TOXIN.			
RX	MEDLINE=11557240; PubMed=11700562; DOI=10.1038/n510198;			
RA	Bradley K.A., Mogridge J., Mourez M., Collier R.J., Young J.A.T.;			
RT	"Identification of the cellular receptor for anthrax toxin.";			
RL	Nature 414:225-229(2001).			
RN	[3]			
RP	SEQUENCE OF 184-564 FROM N.A. (ISOFORM 1), AND SEQUENCE FROM N.A.			
RP	(ISOFORM 3).			
RX	PubMed=14702039; DOI=10.1038/ng1285;			
RA	Ota T., Suzuki Y., Niehikawa T., Otsuki T., Sugiyama T., Irie R.,			
RA	Wakamatsu A., Hayaishi K., Sato H., Nagai K., Kimura K., Makita H.,			
RA	Sekine M., Odayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,			
RA	Yanamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,			
RA	Nagahara K., Murakami K., Yasuda T., Wagaetsuma M.,			
RA	Shiratori A., Sudo H., Hosoi T., Kaku Y., Kodaira H., Kondo H.,			
RA	Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,			
RA	Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,			
RA	Yamazaki M., Nitomiya K., Ishibashi T., Yamashita H., Murakawa K.,			
RA	Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y.,			
RA	Ishida S., Ono Y., Takiguchi S., Watanabe S., Yoshida M., Hotta T.,			
RA	Kuano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,			
RA	Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,			
RA	Iiose N., Mutsaers K., Yuki H., Ohima A., Sasaki N., Aotsuka S.,			
RA	Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,			
RA	Moriya S., Momiyama H., Satoh N., Takami S., Terashima F., Suzuki O.,			
RA	Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,			
RA	Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,			
RA	Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,			
RA	Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujitara T.,			
RA	Ono T., Yamada K., Fujii Y., Ozaki K., Hiro M., Omori Y.,			
RA	Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,			
RA	Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senda T.,			

RA Matsunura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuko Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.,
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.",
RL Nat. Genet. 36:40-45(2004).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 4).
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuller G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rablin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Uedon T.B., Tohiyuki S., Carninci T., Prange C.C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywinski M.I., Skalska J., Smalins D.E.,
RA Schneringer A., Schein J.B., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.",
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP INTERACTION WITH ANTHRAX TOXIN.
RC TISSUE=Placenta;
RX MEDLINE=22608610; PubMed=12700348; DOI=10.1073/pnas.0431098100;
RA Scoble H.M., Rainey G.J.A., Bradley K.A., Young J.A.T.,
RT "Human capillary morphogenesis protein 2 functions as an anthrax toxin
RT receptor.",
RL Proc. Natl. Acad. Sci. U.S.A. 100:5170-5174(2003).
RN [6]
RP SPBCE ISOFORM(S) THAT ARE POTENTIAL NMD TARGET(S).
RX PubMed=14759258; DOI=10.1186/gb-2004-5-2-r8;
RA Hillman R.T., Green R.E., Brenner S.E.,
RT "An unappreciated role for RNA surveillance.",
RL Genome Biol. 5:RESEARCH008.1-RESEARCH008.16(2004).
CC -1- FUNCTION: Cellular role is not yet known.
CC -1- SUBUNIT: Binds to the protective antigen (PA) of Bacillus
CC anthracis. Binding does not occur in the presence of calcium.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Probable).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Comment=Experimental confirmation may be lacking for some
CC isoforms;
CC Name=1;
CC IsoId=Q9H6X2-1; Sequence=Displayed;
CC Note=May be produced at very low levels due to a premature stop
CC codon in the mRNA, leading to nonsense-mediated mRNA decay;
CC Name=2;
CC IsoId=Q9H6X2-2; Sequence=VSP_000444, VSP_000445;
CC Name=3;
CC IsoId=Q9H6X2-3; Sequence=VSP_000446, VSP_000447;
CC Name=4;
CC IsoId=Q9H6X2-4; Sequence=VSP_000448, VSP_000449;
CC -1- TISSUE SPECIFICITY: Highly expressed in tumor endothelial cells
CC but not in normal endothelial cells.
CC -1- DOMAIN: Binding to PA seems to be effected through the VWA domain.
CC -1- SIMILARITY: Belongs to the ATR family.
CC -1- SIMILARITY: Contains 1 VWA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.ebi.ac.uk/announcements/](http://www.ebi.ac.uk/announcements)
CC or send an email to license@ebi-sib.ch).
CC -----
DR EMBL: AF29145; AAK52094.1; -;
DR EMBL: AF421380; AA126496.1; -;
DR EMBL: AK025429; BAB15128.1; ALT_INT.
DR EMBL: AK001463; BAA91707.1; ALT_FRAME.
DR EMBL: BC012074; AAH12074.1; -;
DR GeneW: HGNC:21014; ANTXR1.
DR H-InvDB: HIX0002125; -;
DR MIM: 606410; -;
DR InterPro: IPR008400; Anth_Ig.
DR InterPro: IPR008399; Ant_C.
DR InterPro: IPR02035; VWF_A.
DR Pfam: PF05587; Anth_Ig; 1.
DR Pfam: PF05586; Ant_C; 1.
DR SMART: SM00327; VWA; 1.
DR SMART: PS50234; VWF; 1.
KW Alternative splicing; Glycoprotein; Receptor; Signal; Transmembrane.
FT SIGNAL 1 32
FT CHAIN 33 564
FT DOMAIN 33 321
FT TRANSMEM 322 342
FT DOMAIN 343 564
FT DOMAIN 44 215
FT DOMAIN 360 368
FT DOMAIN 506 564
FT CARBOHYD 166 166
FT CARBOHYD 184 184
FT CARBOHYD 262 262
FT VARSPPLIC 365 368
FT FT
FT VARSPPLIC 369 564
FT FT
FT VARSPPLIC 268 297
FT FT
FT VARSPPLIC 298 564
FT FT
FT VARSPPLIC 319 333
FT FT
FT VARSPPLIC 334 564
FT FT
SQ SEQUENCE 564 AA; 62789 MW; B18A00D5DF2233 CRC64;
Query Match 100.0%; Score 970; DB 1; Length 564;
Best Local Similarity 100.0%; Pred. No. 2,1e-74;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGFDFLFIIDKSGSVLHNNHNIYFVEQLAHKRIISPOLBMSFTVPSRGTTLMKLTEDRE 60
DB 41 GGFDFLFIIDKSGSVLHNNHNIYFVEQLAHKRIISPOLBMSFTVPSRGTTLMKLTEDRE 100
QY QIQGHEELQKLVPGSDTYVHGEFERASEQIYYENRQGYRTASVIALTDGELHEDLFFY 120
DB 101 QIQGHEELQKLVPGSDTYVHGEFERASEQIYYENRQGYRTASVIALTDGELHEDLFFY 160
QY 121 SEERANRSRLGAIIVYCVGVKDNFETQLARIADSKDHVPFVNDGFQALQGIHSILKSC 180
DB 161 SEERANRSRLGAIIVYCVGVKDNFETQLARIADSKDHVPFVNDGFQALQGIHSILKSC 220
QY 181 IETLAIE 187
DB 221 IETLAIE 227
RESULT 2
ATTRI_MOUSE STANDARD; PRT; 562 AA.

CC 09c252-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Anthrax toxin receptor 1 precursor (Tumor endothelial marker 8).
 GN Name=anthrx1; Synonyms=Atx, Tem8;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.
 NC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1)
 RX MEDLINE=21443268; PubMed=11559528;
 RA Carson-Walter E.B., Watkins D.N., Nanda A., Vogelstein B.,
 RT Kinzler K.W., St Croix B.
 RT "Cell surface tumor endothelial markers are conserved in mice and
 humans."
 RL Cancer Res. 61:6649-6655 (2001).
 RN [2]
 RP SEQUENCE OF 88-562 FROM N.A. (ISOFORM 2).
 RC STRAIN=C57BL/6J; TISSUE=Embryo;
 RX MEDLINE=23354683; PubMed=12466651; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kanakawa T., Adachi Y., Bono H., Kondo S.,
 RA Nikaido I., Otsu N., Saito R., Suzuki H., Yamahata I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gotohori T.,
 RA Baldairelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schmitt L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Chochia C., Corbani L.E., Cousins S.,
 RA Dalle E., Dragani T.A., Fletcher C.F., Forrest A., Frerz K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Guelinich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzielski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lemhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Milt H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Petrea G., Picot G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sanderlin A., Schneider C., Sempke C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilmig L.G., Wyrshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yun Z., Zavolan M., Zhu Y., Zimmer A., Carlini P., Hayatsu N.,
 RA Hirotsuka-Ishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
 RA Shirozaki T., Waki K., Kawai J., Aizawa K., Arikawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.,
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 CC -1- FUNCTION: Cellular role is not yet known. (PA) of Bacillus
 anthracis (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Probable).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q9CZ52-1; Sequence=displayed;
 CC Name=2;
 CC IsoId=Q9CZ52-2; Sequence=VSP 000450;
 CC Note=No experimental confirmation available;
 CC -1- DOMAIN: Binding to PA seems to be effected through the VWA domain
 (By similarity).
 CC -1- SIMILARITY: Belongs to the ATR family.
 CC -1- SIMILARITY: Contains 1 VWA domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>)

CC		or send an email to license@lab-sib.ch).
DR	EMBL; AF378762; AAL1999.1;	-
DR	EMBL; AK013005; BAB28591.1; ALT_INIT.	
DR	MGI; MGI:1916788; Antxr1.	
DR	InterPro; IPR008400; Antch_Ig.	
DR	InterPro; IPR008359; Ant_C.	
DR	InterPro; IPR02035; VWF_A.	
DR	pfam; PF05587; Antch_Ig; 1.	
DR	pfam; PF05586; Ant_C; 1.	
DR	pfam; PF00092; VWF; 1.	
KM	PROSITE; PS50234; VWF.A; 1.	
FT	Alternative splicing; Glycoprotein; Receptor; Signal; Transmembrane.	
FT	SIGNAL	1
FT	CHAIN	30
FT	POTENTIAL	1
FT	DOMAIN	31 562
FT	TRANSMEM	31 319
FT	POTENTIAL	320 340
FT	CYCLOPLASMIC	341 562
FT	DOMAIN	42 213
FT	DOMAIN	358 366
FT	DOMAIN	501 562
FT	CARBOHYD	164 164
FT	CARBOHYD	182 182
FT	CARBOHYD	260 260
FT	VARSPLIC	477 562
FT	GRCINPFRVKNKSQPAKYPLNNTYHPSSPPAPITYTTPPEPAP	
FT	HCPPEPAPSPAPPPIPSPPSTLPPOAPPPRPARPPSPAPP	
FT	RPSV -> RFRGMRLTTCIGSKVHYPGHDXGPETPLLKQA	
FT	WMFSSPLERARQ (in isoform 2).	
FT	/PTTD-USP 000450	
FT	SEQUENCE	562 AA; 62308 MW; 6AC92049B4BB4F7C CRC64;
Query Match	Best Local Similarity	99.1%; Score 961; DB 1; Length 562;
Matches	Conservative	185; Indels 1; Mismatches 1; Gaps 0;
QY	1 GGPDLYFLDKSGSVLHHANEIYFVEQLAKKFISPOLRNSFIYFSTRGTIMKLTEBRE	60
DB	39 GGFDLYFLDKSGSVLHHANEIYFVEQLAHRFISPOLRMSEIFIVSTRGTIMKLTEBRE	98
QY	61 QIRGLAEVLQVLRGGDTVMHEGEFRASEOIYYENRGQYRASVIYIALTOGELHEDLFY	120
DB	99 QIRGLAEVLQVLRGGDTVMHEGEFRASEOIYYENRGQYRASVIYIALTOGELHEDLFY	158
QY	121 SERANRSRDIGAIVYCVGVDPFNETOLARIADSKDHVPFNDFOLFALOGIHSLKKSC	180
DB	159 SERANRSRDIGAIVYCVGVDPFNETOLARIADSKDHVPFNDFOLFALOGIHSLKKSC	218
QY	181 IEILAAE 187	
DB	219 IEILAAE 225	
RESULT 3		
ID	ATR2_HUMAN	STANDARD; PRT; 489 AA.
AC	PS8335; O86U11; O8NB13; O96NC7;	
DT	28-FEB-2003 (Rel. 41, Created)	
DT	10-OCT-2003 (Rel. 42, Last sequence update)	
DT	05-JUL-2004 (Rel. 44, Last annotation update)	
DE	Antitran toxin receptor 2 precursor (Capillary morphogenesis protein-2) (CMG-2).	
DS	Name=ANTXR2; Synonyms=CMG2;	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX	NCBI_TaxID=9606;	
RN	SEQUENCE FROM N.A. (ISOFORM 2).	
RX	MEDLINE=21539596; PubMed=11683410;	
RA	Bell S.E., Mavila A., Salazar R., Bayless K.J., Kanagala S.,	
RA	Maxwell S.A., Davis G.B.;	
RT	"Differential gene expression during capillary morphogenesis in 3D collagen matrices: regulated expression of genes involved in basement	

RT membrane matrix assembly, cell cycle progression, cellular
 RT differentiation and G-protein signaling.";
 RL J. Cell Sci. 114:2755-2773(2001).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND INTERACTION WITH ANTHRAX TOXIN.
 RC TISSUE=Placenta;
 RX MEDLINE=22608610; PubMed=12700348; DOI=10.1073/pnas.0431098100;
 RA Scobie H.M., Rahney G.J.A., Bradley K.A., Young J.A.T.;
 RT "Human capillary morphogenesis protein 2 functions as an anthrax toxin
 receptor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:5170-5174(2003).
 RN [3]
 RP SEQUENCE OF 78-489 FROM N.A. (ISOFORM 3), AND SEQUENCE FROM N.A.
 RA (ISOFORM 4).
 RC TISSUE=synovial cell;
 RX PubMed=14702039; DOI=10.1038/ng1285;
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
 Nagahara K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
 Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
 Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
 Omura Y., Abe K., Kamihara K., Katsura N., Sato K., Tanikawa M.,
 Yamazaki M., Nimomiya K., Ishibashi T., Yamashita H., Murakawa K.,
 Fujimori K., Tanai H., Kimata M., Watanabe M., Hiwaka S., Chiba Y.,
 Iehida S., Oho Y., Takiguchi S., Watanabe S., Yoshida M., Hoshida T.,
 Kusanaga Y., Togeishi S., Komai F., Hara R., Takeuchi K., Arita M.,
 Imose N., Mueshishima K., Yuzuki H., Ooshima A., Sasaki N., Aoshima S.,
 Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,
 Moriya S., Nomiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 Hishiyaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
 Oho T., Yamada K., Fujii Y., Ozaki K., Hiroo K., Ohmori Y.,
 Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
 Matsunura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 Okumura K., Nagase T., Nomura N., Kikuchi H., Maeno Y., Yamashita R.,
 Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
 RT "Complete sequencing and characterization of 21,243 full-length human
 cDNAs.";
 RL Nat. Genet. 36:40-45(2004).
 CC -1- FUNCTION: Cellular role is not yet known.
 CC -1- SUBUNIT: Binds to the protective antigen (PA) of Bacillus
 anthracis in a divalent cation-dependent manner, with the
 following preference: calcium > manganese > magnesium > zinc.
 CC Seems to bind to collagen type IV and laminin.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (isoforms 1 and 2).
 CC Secreted (isoform 3). Isoform 1 is expressed at the cell surface
 while isoform 2 is predominantly expressed within the endoplasmic
 reticulum and not at the plasma membrane.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=4;
 CC Name=1;
 CC IsoId=P58335-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=P58335-2; Sequence=VSP_008343;
 CC Name=3;
 CC IsoId=P58335-3; Sequence=VSP_008344, VSP_008345;
 CC Note=No experimental confirmation available;
 CC Name=4;
 CC IsoId=P58335-4; Sequence=VSP_008346;
 CC Note=No experimental confirmation available;
 CC -1- TISSUE SPECIFICITY: Expressed in colon, heart, kidney, lung,
 liver, peripheral blood leukocytes, placenta, skeletal muscle,
 small intestine and spleen.
 CC -1- DOMAIN: Binding to PA seems to be effected through the VWA domain.
 CC -1- SIMILARITY: Belongs to the ATR family.

CC -1- SIMILARITY: Contains 1 VWA domain.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AY040326; AAK7222.1; -;
 DR EMBL; AY233452; AAP04016.1; -;
 DR EMBL; AK055636; BAB70976.1; ALT_INT.
 DR EMBL; AK091721; BAC03731.1; -;
 DR Genew; HGNC:21732; ANTXR2.
 DR MIM; 608041; -;
 DR InterPro; IPRO08399; Ant_C.
 DR InterPro; IPRO02035; VWF_A.
 DR Pfam; PF00586; Ant_C_1.
 DR Pfam; PF00092; VWA_1.
 DR PROSITE; PS50234; VWF_1.
 KW Alternative splicing; Glycoprotein; Receptor; Signal; Transmembrane.
 FT SIGNAL 1
 FT CHAIN 33
 FT DOMAIN 34
 FT TRANSMEM 34
 FT DOMAIN 319
 FT DOMAIN 342
 FT DOMAIN 44
 FT CARBOHYD 250
 FT CARBOHYD 260
 FT VARSPLIC 213
 FT VARSPLIC 290
 FT VARSPLIC 322
 FT VARSPLIC 323
 FT VARSPLIC 477
 FT VARSPLIC 489
 FT CONFLICT 357
 FT SEQUENCE 489 AA; 53692 MW; B9F679DB7586EB27 CRC64;
 SQ
 Query Match 357 357 56.7%; Score 550; DB 1; Length 489;
 Best Local Similarity 59.9%; Pred. No. 1.3e-38;
 Matches 109; Conservative 35; Mismatches 36; Indels 2; Gaps 1;
 QY 3 FDIYFLDKSGSYLHMNIEYVEQLAKKFIPOALRMSTIVSTRGTLMLKLTDEBQI 62
 DB 43 FDIYFVLDKSGSYVANNMIEYVFOQLAEKRFVPEMKLSLIVFSSQATITLPLTGDRGI 102
 QY 63 ROGLEELQKLPAGDPTVMEHGFASRBOIYENRGKRTASVIALTDEGHEDLFYSE 122
 DB 103 SKLEEDLAKRSPVGERIYHGLKLANEQI-QVAGLKTSIIITLTDGLDGLVPSYAE 160
 QY 123 REANRSRLGAIYCVGVDPNETOLARLADSKOHVFPVNDGFOALGGIHSILKSCIE 182
 DB 161 KEAKISRSIGASYVCVGLDFEQALERLADSKOEQVFPVNGQALKGIIINSIAGCTE 220
 QY 183 IL 184
 DB 221 IL 222
 RESULT 4
 ID Q6DFX2 PRELIMINARY; PRT; 487 AA.
 AC Q6DFX2;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, last annotation update)
 DE Anthrax toxin receptor 2.

GN Name=Antxr2;
 OS Mus musculus (mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 [1]
 RN
 RP
 RC
 RA STRAIN=C57BL/6; TISSUE=Eye;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Krauner R.D., Collins F.S., Wagner K.H., Sherman C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkin R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marisica K., Farmer A.A., Rubin G.M., Hong L.,
 RA Slatopolsky M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toshitoki S., Carninci P., Prange C.,
 RA Rhee S.S., Loguelfano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Boeck S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunnaracne P.H.,
 RA Richards S., Wotley K.C., Hale S., Garcia A.M., Gay L.U., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fehey J., Helton E., Kesteman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywnicki M.I., Skelton U., Small D.E., Schermer A., Schein J.E.,
 RA Jones S.J., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RL
 RN
 RP
 RC
 RA STRAIN=C57BL/6; TISSUE=Eye;
 RA Strauberg R.,
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC076595; AAT76555.1; -
 DR GO: GO:0016021; C: integral to membrane; IEA.
 DR GO: GO:0004872; P: receptor activity; IEA.
 DR InterPro: IPR008400; Antc Ig.
 DR InterPro: IPR008399; Antc Ig.
 DR InterPro: IPR002035; VWF A.
 DR Pfam: PF05587; Antc Ig 1.
 DR Pfam: PF05386; Antc Ig 1.
 DR Pfam: PF00092; VWA 1.
 DR SMART: SM00327; VWA 1.
 DR PROSITE: PSS0234; VWFA 1.
 KM Receptor.
 SO SEQUENCE 487 AA; 53184 MW; 61A400D60BC8D69 CRC64;
 Query Match 56.5%; Score 548; DB 2; Length 487;
 Best Local Similarity 59.9%; Pred. No. 2e-38;
 Matches 109; Conservative 32; Mismatches 39; Indels 2; Gaps 1;
 QY 3 PDLYFIIDKSGSVLAHNNIYFYEQIAHKFISQILMSFIVSTRTGTLTKLTEDREOI 62
 DB 43 PDLYFVLDKSGSVANNMIIEYFNHQLTERFVSPKMLSPFVSSQATILPLTGDRYKI 102
 QY 63 ROGIEELQKVLPGSDTYMHGSPERASQIYYBNQGRITASVILLADGELHEDLFEYSE 122
 DB 103 GKGIEDKAVKPVGETIYHRLKLANEOT--QNNAGSKASTIITLTDGKLDGLVPSYAE 160
 QY 123 REANRSHDLAIVYCVGVGFNETQTLARIADSKDHPVPVNDGFQALOGIIHSLKSCIE 182
 DB 161 NNAKSSSLASVYCVGVLPFEGQQLERIDSKQGVFPVPGKPGALGIIINSILAQCTE 220
 QY 183 IL 184
 DB 221 IL 222
 RESULT 5
 Q8BVM2 PRELIMINARY; PRT; 641 AA.
 AC Q8BVM2;
 ID Q8BVM2;
 GO: GO:0016021; C: integral to membrane; IEA.

DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Mus musculus adult male testis cDNA, RIKEN full-length enriched
 DE library, clone 493343011 product: hypothetical Prolin-rich region/von
 DE Willebrand factor type A domain containing protein, full insert
 DE sequence.
 OS Mus musculus (mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 [1]
 RN
 RP
 RC
 RA STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.,
 RT "High-efficiency full-length cDNA cloning";
 RT Meth. Enzymol. 303:19-44(1999).
 RL
 RN
 RP
 RC
 RA STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA RIKEN PANTOM Consortium,
 RT "Functional annotation of a full-length mouse cDNA collection";
 RT Nature 409:685-690(2001).
 RL
 RN
 RP
 RC
 RA STRAIN=C57BL/6J; TISSUE=Testis;
 RA The PANTOM Consortium,
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs";
 RT Nature 420:563-573(2002).
 RL
 RN
 RP
 RC
 RA STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=20493374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.,
 RT "Normalization and subcloning of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes";
 RT Genome Res. 10:1617-1630(2000).
 RL
 RN
 RP
 RC
 RA STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Kono H., Akiyama J., Nishi K., Kitzunai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaueuchi S., Ikegami T., Kaishiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Matshiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai U.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer";
 RT Genome Res. 10:1757-1771(2000).
 RL
 RN
 RP
 RC
 RA STRAIN=C57BL/6J; TISSUE=Testis;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hatanaka T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
 RA Horii F., Imotani K., Ishii Y., Itoh M., Kagaawa I., Kasukawa T.,
 RA Katoh H., Kawai U., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohnato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takekida Y., Tanaka T.,
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK077206; BAC36683.1; -
 GO: GO:0016021; C: integral to membrane; IEA.

```

DR GO:0004872; F:receptor activity; IEA.
DR InterPro: IPR008400; Antch_1g.
DR InterPro: IPR002035; VWF_A.
DR Pfam: PF05587; Antch_1g; 1.
DR Pfam: PF00092; VMA; 1.
DR PRINTS: PRO0453; VWFADOMAIN.
DR SMART: SM00327; VMA; 1.
DR PROSITE: PS50234; VWF_A; 1.
DR Hypothetical protein.
SQ SEQUENCE 641 AA; 70415 MW; 199E300730BC85E3 CRC64;

Query March 40.0%; Score 388; DB 2; Length 641;
Best Local Similarity 43.9%; Pred. No. 1.3e-24;
Matches 82; Conservative 33; Mismatches 72; Indels 0; Gaps 0;

QY 1 GGFDPYILDKSGSVLHHMNIYYFVQLAHKFIISPOLRMSFIYFSTRGTTMLKLTEDRE 60
DB 73 GIFPLVILVLDKSGSVADNWHIYSEFAGLVKFTNPMLRISIIYSTEAVILPLDTSK 132
QY 61 QIRGSEBELQKVLPGSGDTYHMEGFERASEQIYENRGYRTASVITALTDELHEDLPFY 120
DB 133 EIKNSLLVLSIVSGVLTTHQKGRKANEOIRKSTLGRIVNSVITALTDELHEDLPFY 192
QY 121 SEREANSRDLGALIVCGVDFMETQLARIADSKDHVPVNDGFOALOGIHSILKSK 180
DB 193 TMEAKAKRRMGALIVTVGVPMYSKQOLVNIAGDPDRCFGVDEGFSALGVDPLETSK 252
QY 181 IEILAAE 187
DB 253 TELLSVQ 259

RESULT 6
Q9BP08 PRELIMINARY; PRT; 1332 AA.
ID Q9BP08
AC Q9BP08;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Integrin alpha Hrl precursor.
GN Name=HrTcA1.
OS Halocynthia roretzi (Sea squirt).
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Scolidobranchia; Pyridae; Halocynthia.
OX NCBI_TaxID=7729;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hemocyte;
RX MEDLINE=21103187; PubMed=11160215;
RA Miyazawa S., Azumi K., Nonaka M.,
RT "Cloning and characterization of integrin alpha subunits from the
RT solitary ascidian, Halocynthia roretzi.";
RL J. Immunol. 166:1710-1715 (2001).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the Integrin alpha chain family.
DR EMBL: AB048261; BAB21479.1; -.
DR HSSP: P11215; 1BHQ.
DR GO: GO:0016021; C:Integral to membrane; IEA.
DR GO: GO:0008305; C:Integrin complex; IEA.
DR GO: GO:0005515; F:Protein binding; IEA.
DR GO: GO:0007160; P:Cell-matrix adhesion; IEA.
DR GO: GO:0007229; P:Integrin-mediated signaling pathway; IEA.
DR InterPro: IPR000413; Integrin_alpha.
DR Pfam: PF00092; VMA; 1.
DR PRINTS: PRO1185; INTEGRINA.
DR SMART: SM00191; Int_alpha; 5.
DR SMART: SM00327; VMA; 1.
DR PROSITE: PS50234; VWF_A; 1.
KM Cell adhesion; Integrin; Signal; Transmembrane.
FT SIGNAL 1 Potential.
FT CHAIN 31 1332 Integrin alpha Hrl.

```

```

SQ SEQUENCE 1332 AA; 145851 MW; 0D9108D2B05CFEAE CRC64;

Query Match 15.1%; Score 146; DB 2; Length 1332;
Best Local Similarity 24.5%; Pred. No. 0.0014;
Matches 52; Conservative 34; Mismatches 60; Indels 66; Gaps 7;

QY 2 GFDLYFLDKSGSVLHHMNIYYFVQLAHKFIISPOLRMSFIYFSTRGTTMLKLTEDRE 61
DB 204 GAVVLFPVLDGSGSVGNKPFVDKQVKNVT-----AKLDIGKEI 241
QY 62 IROGLEBELQKVLPG-----GDTYHMEGFERASEQIYENRGYRT----- 101
DB 242 VRGVGVQYSHYVAGSKINQKYITTHESISGEFLNDFENAVDRI---QLQYTTTYGRA 298
QY 102 -----ASVITALTDELHEDLPFYSEREANSRDLGALIVCGVDFNET 146
DB 299 LQKIRDFDDAVYIGNKQVLLTLTDGQAKDKILP--NANRLNKGIATPAVGGEYDIS 356
QY 147 QLARIA---DSKDHVPVNDGFOALOGIHSI 175
DB 357 ELKLASGTDSTURVFTVD-FGELDSIVKSL 387

RESULT 7
ITAD RAT
ID ITAD RAT
AC Q90YE7;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Integrin alpha-D precursor.
GN Name=Itgad;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA O'Brien M.M., VanderVlieten M., Kilgannon P.D., Dietrich G.,
RA Gallatin W.M.;
RT "Cloning of rat alpha D, a novel beta 2 integrin.";
RL Submitted (MIG-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Integrin alpha-D/beta-2 is a receptor for ICAM3 and
CC VCAM1. May play a role in the atherosclerotic process such as
CC clearing lipoproteins from plaques and in phagocytosis of blood-
CC borne pathogens, particulate matter, and senescent erythrocytes
CC from the blood (By similarity).
CC -1- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-D
CC associates with beta-2 (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -1- DOMAIN: The integrin I-domain (insert) is a VMA domain. Integrins
CC with I-domains do not undergo protease cleavage.
CC -1- SIMILARITY: Belongs to the Integrin alpha chain family.
CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
CC -1- SIMILARITY: Contains 1 VMA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AF021334; AAF21241.1; -.
DR HSSP: P11215; 1BHQ.
DR InterPro: IPR000413; Integrin_alpha.
DR InterPro: IPR002035; VWF_A.
DR Pfam: PF01839; FG-GAP; 3.
DR Pfam: PF00357; Integrin_alpha; 1.
DR Pfam: PF00092; VMA; 1.
DR PRINTS: PRO1185; INTEGRINA.

```



```

DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VMA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VMA; 1.
KM Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
KW Repeat; Signal; Transmembrane.
FT SIGNAL 1 19 Potential.
FT CHAIN 20 1161 Integrin alpha-D.
FT DOMAIN 20 1100 Extracellular (Potential).
FT TRANSMEM 1101 1121 Cytoplasmic.
FT DOMAIN 1122 1161 Potential.
FT REPEAT 34 87 FG-GAP 1.
FT REPEAT 88 88 FG-GAP 2.
FT REPEAT 152 334 VMA.
FT REPEAT 352 402 FG-GAP 3.
FT REPEAT 403 454 FG-GAP 4.
FT REPEAT 456 517 FG-GAP 5.
FT REPEAT 519 577 FG-GAP 6.
FT REPEAT 582 634 FG-GAP 7.
FT CA_BIND 467 475 Potential.
FT CA_BIND 531 539 Potential.
FT CA_BIND 594 602 Potential.
FT SITE 1126 1130 GPCR motif.
FT DISULFID 69 76 By similarity.
FT DISULFID 108 126 By similarity.
FT DISULFID 656 711 By similarity.
FT DISULFID 769 775 By similarity.
FT DISULFID 845 860 By similarity.
FT DISULFID 993 1017 By similarity.
FT DISULFID 1022 1027 By similarity.
FT CARBOHYD 61 61 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 89 89 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 245 245 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 393 393 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 696 696 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 734 734 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 784 784 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 907 907 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 936 936 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1045 1045 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 1161 AA; 126600 MW; 2258491A984A705E CRC64;

Query Match 14.6%; Score 141.5; DB 1; Length 1161;
Best Local Similarity 27.7%; Pred. No. 0.0029;
Matches 53; Conservative 38; Mismatches 75; Indels 25; Gaps 10;

QY 4 DLYPIIDKSGSV-LHHNNEIYFVEQLAHKFIISPOLHMSPIVSTRTGTTLMKTE----- 57
DB 152 DIAFLIDGSGSINORDPAQMDPFKALMGEPASTSTLFSLMQYSNLIKHTFTFEPKNIL 211
QY 58 DREIRGRLLEQLVLCGDTYMHGEPERASEQIYYENRCQYRA-SVIALTDTGELHED 116
DB 212 DPGSLVDPIVQLQ-----GLIYATYATGIRTYWELFHSKNGSRKSAKKILLVITDGQRYD 266
QY 117 LFPYSR--REANRSRDIGAIYVCVGVQ-FNE-TOLARI-----ADSKDHVPVNDGFOA 167
DB 267 PLEYSDVIYPADKX---GIIRYALGVGDARFEPALKEINTIGSAPPDHVFYKGN-FAA 322
QY 168 LOGIHSILKK 178
DB 323 LRSIORQLQEK 333

RESULT 8
Q8T6U5 PRELIMINARY; PRT; 441 AA.
AC Q8T6U5;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Proximal thread matrix protein 1 variant a.
OS Mytilus edulis (Blue mussel).

```

```

OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_TaxID=6550;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22313036; PubMed=12425661; DOI=10.1021/bm0255903;
RA Sun C., Lucas J.M., Waite J.H.;
RT "Collagen-binding matrix proteins from elastomeric extraorganicmic
RT byssal fibers.";
RL Biomacromolecules 3:1240-1248 (2002).
DR EMBL; AF414454; AAL83537.1;
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF00092; VMA; 2.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00327; VMA; 2.
DR PROSITE; PS50234; VMA; 2.
KW Matrix protein.
SQ SEQUENCE 441 AA; 47543 MW; 881DBBD36B891D2B CRC64;

Query Match 14.5%; Score 141; DB 2; Length 441;
Best Local Similarity 25.3%; Pred. No. 0.001;
Matches 50; Conservative 46; Mismatches 76; Indels 26; Gaps 12;

QY 1 GGFPIYIIDKSGSV-LHHNNEIYFVEQLAHKFIISPOLHMSPIVSTRTGTTLM 53
DB 238 GHADIAFVFPASSISNNNNNNNNGLMKDFMKDIYDRNKTKGPDGTQPAVTVFADRAATKQF 297
QY 54 KITE--DREIRGRLLEQLVLPG--GDTYMHGEPERASEQIYYENRCQY--YTAIVII 106
DB 298 GLNDYSSKAIKKAID--KVTPEITIGQTAIGDELNARLEV-FPNNGGREGVEQKVI 353
QY 107 ALTFDEL--HEDLFYSEERANRSRDIGAIYVCVGV-KDENEIOLARIADSKDHVPVND 163
DB 354 LITDGNNGHKS-----PEHSSSLRKEGVIVVAIGVGTGLKSLINIASSEEVF-TTS 408
QY 164 GPQALQGIHSILKKSCI 181
DB 409 SFNKLSKIMENVVTLACM 426

RESULT 9
Q8T5C3 PRELIMINARY; PRT; 444 AA.
AC Q8T5C3;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Proximal thread matrix protein 1b.
OS Mytilus edulis (Blue mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_TaxID=6550;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22313036; PubMed=12425661; DOI=10.1021/bm0255903;
RA Sun C., Lucas J.M., Waite J.H.;
RT "Collagen-binding matrix proteins from elastomeric extraorganicmic
RT byssal fibers.";
RL Biomacromolecules 3:1240-1248 (2002).
DR EMBL; AY053390; AAL17973.1;
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF00092; VMA; 2.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00327; VMA; 2.
DR PROSITE; PS50234; VMA; 2.
KW Matrix protein.
SQ SEQUENCE 444 AA; 47815 MW; D2C605347450C931 CRC64;

Query Match 14.5%; Score 141; DB 2; Length 444;
Best Local Similarity 25.3%; Pred. No. 0.001;
Matches 50; Conservative 46; Mismatches 76; Indels 26; Gaps 12;

```

```
QY 1 GGFLLYITLXKSGV-----LHNNETIYFVQLAHKF--ISPO-LRMSPIVSTRTGTTLM 53
D 241 GHADIAEVFASSSSINANNPNNGYGLMKDFMDYDRFKTKGPPDDOTQPAVWYFADRAKQF 300
QY 54 KLTE--REOQIRQGLLELQKVLPG--GDYWHGEFEPASQIYYENNQG---YNTASVYI 106
D 301 GLKDYSSKAEIKKAIID--KTPPSITIGTATIGDLENNRLEV--FPNNNGGSEEVQKQVI 356
QY 107 ALTDGEL--HEDLFYSREANRSRDIGAIYCVGV-KDPNETQIARIADSKDHVFPVND 163
D 357 LITDGNNGHKS-----PEHSSLLRKSGVVAIGVGGLSKSLINLASEEVVF--TTS 411
QY 164 GFQALQGIHSHILKKSCT 181
D 412 SFNNLSKIMENVYLAQM 429

RESULT 10
CALC_MOUSE STANDARD; PRT; 3119 AA.
ID Q60847; P70322;
DT 15-JUL-1998 (Rel. 36, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Collagen alpha 1(XII) chain precursor.
GN Name=Coll2a1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING (ISOFORMS XIIA-1 AND
RP XIIA-2).
RC STRAIN=C57BL/6J; TISSUE=Skin fibroblast;
RX MEDLINE=99348349; PubMed=10419533; DOI=10.1074/jbc.274.31.22053;
RA Kania A.M., Reichenberger E., Baur S.T., Karimbux N.Y., Taylor R.W.,
RA Olsen B.R., Nishimura I.;
RT "Structural variation of type XII collagen at its carboxyl-terminal
RT NCI domain generated by tissue-specific alternative splicing.";
RL J. Biol. Chem. 274:22053-22059(1999).
CC -1- FUNCTION: Type XII collagen interacts with type I collagen-
CC containing fibrils, the COL1 domain could be associated with the
CC surface of the fibrils, and the COL2 and NC3 domains may be
CC localized in the pericellular matrix (By similarity).
CC -1- SUBUNIT: Trimer of identical chains each containing 190 kDa of
CC nontriple-helical sequences (By similarity).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Comment=The final tissue form of collagen XII may contain
CC homotrimers or any combination of the various isoforms;
CC Name=XIIA-1;
CC IsoId=Q60847-1; Sequence=Displayed;
CC Name=XIIA-2; Synonyms=ER#K;
CC IsoId=Q60847-2; Sequence=VSP_001151, VSP_001152;
CC Name=XIIB-1;
CC IsoId=Q60847-3; Sequence=VSP_001150;
CC Name=XIIB-2;
CC IsoId=Q60847-4; Sequence=VSP_001150, VSP_001151, VSP_001152;
CC -1- TISSUE SPECIFICITY: Highest expression in tendons, perichondrium,
CC skin, cornea, sclera, blood vessels, and pericardium.
CC -1- DEVELOPMENTAL STAGE: The long NC3 XIIA isoforms are predominant at
CC early stages (ED7 and 11); at later stages of development (ED15
```

```
CC and 17) the short NC3 XIIB forms become the major forms. As the
CC short NC3 forms become the major product, the long splice variant
CC continues to be expressed in several tissues, even after birth.
CC The long NCI isoforms, XIIA-1 and XIIB-1, peak in 15-day old
CC embryos and decrease in 17-day old ones. The expression of the
CC short NCI form XIIB-2 remains constant throughout late stages of
CC embryonic development (ED15 and ED17).
CC -1- PTM: The triple-helical tail is stabilized by disulfide bonds at
CC each end (By similarity).
CC -1- PTM: Prolines at the tripeptide position of the tripeptide repeating
CC unit (G-X-Y) are hydroxylated in some or all of the chains (By
CC similarity).
CC -1- PTM: O-glycosylation of isoform XIIA-2; glycosaminoglycan of
CC chondroitin-sulfate type (By similarity).
CC -1- SIMILARITY: Belongs to the fibril-associated collagens with
CC interrupted helices (FACIT) family.
CC -1- SIMILARITY: Contains 18 fibronectin type III domains.
CC -1- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -1- SIMILARITY: Contains 4 WFPA domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U25652; AAA99719.1; ALT_SEQ.
DR EMBL; U57095; AAB07047.1; -.
DR HSSP; P18614; LMHP.
DR MGD; MGI:88448; Col12a1.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR008985; Cona1 like_1ec_g1.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR006957; FN_III-like.
DR InterPro; IPR003129; TSP_N.
DR InterPro; IPR002035; WFPA.
DR Pfam; PF01391; Collagen; 5.
DR Pfam; PF02210; TSP_N; 1.
DR Pfam; PF00092; WFPA; 4.
DR PRINTS; PR00453; WFPA DOMAIN.
DR SMART; SM00060; FN3; 18.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00327; VWA; 4.
DR PROSITE; PS50853; FN3; 18.
DR PROSITE; PS50234; WFPA; 4.
KW Alternative splicing; Cell adhesion; Collagen; Extracellular matrix;
KW Glycoprotein; Hydroxylation; Repeat; Signal; Structural protein.
FT SIGNL 1 24
FT CHAIN 25 3119
FT DOMAIN 25 112 Fibronection type-III 1.
FT DOMAIN 140 316 WFPA 1.
FT DOMAIN 333 422 Fibronection type-III 2.
FT DOMAIN 444 620 WFPA 2.
FT DOMAIN 635 723 Fibronection type-III 3.
FT DOMAIN 726 814 Fibronection type-III 4.
FT DOMAIN 817 905 Fibronection type-III 5.
FT DOMAIN 908 997 Fibronection type-III 6.
FT DOMAIN 999 1087 Fibronection type-III 7.
FT DOMAIN 1090 1179 Fibronection type-III 8.
FT DOMAIN 1203 1375 WFPA 3.
FT DOMAIN 1388 1476 Fibronection type-III 9.
FT DOMAIN 1478 1567 Fibronection type-III 10.
FT DOMAIN 1569 1656 Fibronection type-III 11.
FT DOMAIN 1660 1747 Fibronection type-III 12.
FT DOMAIN 1758 1847 Fibronection type-III 13.
FT DOMAIN 1849 1937 Fibronection type-III 14.
FT DOMAIN 1939 2028 Fibronection type-III 15.
FT DOMAIN 2030 2119 Fibronection type-III 16.
FT DOMAIN 2121 2208 Fibronection type-III 17.
FT DOMAIN 2212 2298 Fibronection type-III 18.
```

```
FT DOMAIN 2329 2501 WMA 4
FT DOMAIN 2525 2717 TSP N-terminal.
FT DOMAIN 2456 2751 Nonhelical region (NC3).
FT DOMAIN 2752 2899 Triple-helical region (COL2) with 1
FT DOMAIN 2900 2942 Triple-helical region (NC2).
FT DOMAIN 2943 3045 Triple-helical region (COL1) with 2
FT DOMAIN 3046 3119 Nonhelical region (NC1).
FT SITE 866 868 Cell attachment site (Potential).
FT SITE 2784 2786 Cell attachment site (Potential).
FT SITE 2896 2898 Cell attachment site (Potential).
FT MOD_RES 2945 2945 Hydroxyproline (By similarity).
FT MOD_RES 2948 2948 Hydroxyproline (By similarity).
FT MOD_RES 2951 2951 Hydroxyproline (By similarity).
FT MOD_RES 2960 2960 Hydroxyproline (By similarity).
FT MOD_RES 2966 2966 Hydroxyproline (By similarity).
FT MOD_RES 2969 2969 Hydroxyproline (By similarity).
FT MOD_RES 2972 2972 Hydroxyproline (By similarity).
FT MOD_RES 2984 2984 Hydroxyproline (By similarity).
FT MOD_RES 3001 3001 Hydroxyproline (By similarity).
FT MOD_RES 3004 3004 Hydroxyproline (By similarity).
FT MOD_RES 3015 3015 Hydroxyproline (By similarity).
FT MOD_RES 3024 3024 Hydroxyproline (By similarity).
FT MOD_RES 3027 3027 Hydroxyproline (By similarity).
FT MOD_RES 3030 3030 Hydroxyproline (By similarity).
FT DOMAIN 869 872 Poly-Thr.
FT CARBOHYD 704 704 O-linked (GlcNAc-. ) (Potential).
FT CARBOHYD 802 802 O-linked (Xyl-. ) (Chondroitin sulfate)
FT CARBOHYD 893 893 O-linked (Xyl-. ) (Chondroitin sulfate)
FT CARBOHYD 985 985 O-linked (Xyl-. ) (Chondroitin sulfate)
FT CARBOHYD 1769 1769 N-linked (GlcNAc-. ) (Potential).
FT CARBOHYD 2212 2212 N-linked (GlcNAc-. ) (Potential).
FT CARBOHYD 2533 2533 N-linked (GlcNAc-. ) (Potential).
FT CARBOHYD 2684 2684 N-linked (GlcNAc-. ) (Potential).
FT VARSPLIC 25 1190 Missing (in isoform XIIB-1 and isoform
FT XIIB-2).
FT VARSPLIC 3062 3064 /FTId=VSP_001150.
FT VARSPLIC 3065 3119 Epy -> GSG (in isoform XIIB-2 and isoform
FT XIIB-2).
FT VARSPLIC 3065 3119 /FTId=VSP_001151.
FT VARSPLIC 3065 3119 Missing (in isoform XIIB-2 and isoform
FT XIIB-2).
SQ SEQUENCE 3119 AA, 340239 MW, 981F999C86AB3251 CRC64,
Query Match 14.3%; Score 139; DB 1; Length 3119;
Best Local Similarity 30.3%; Pred. No. 0.015;
Matches 60; Conservative 37; Mismatches 75; Indels 26; Gaps 12;
```

```
AC Q8T5C2;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Proximal thread matrix protein 1.
OS Mytilus galloprovincialis (Mediterranean mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Periomorphia; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_TaxID=29158;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2231036; PubMed=12425661; DOI=10.1021/bm0255903;
RA Sun C., Lucas J.M., Waite J.H.;
RT "Collagen-binding matrix proteins from elastomeric extraorganic
RT byssal fibers.";
RL Biomacromolecules 3:1240-1248 (2002).
DR EMBL; AY053391; AL117974.1; -.
DR HSSP; P20701; IMUN.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF00092; VMA; 2.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00327; VMA; 2.
DR PROSITE; PS50234; WMA; 2.
KW Matrix protein.
SQ SEQUENCE 453 AA; 48784 MW; D60497F5C0C51BED CRC64;
Query Match 14.2%; Score 138; DB 2; Length 453;
Best Local Similarity 25.3%; Pred. No. 0.0019;
Matches 50; Conservative 44; Mismatches 78; Indels 26; Gaps 12;
```

```
FT DOMAIN 2329 2501 WMA 4
FT DOMAIN 2525 2717 TSP N-terminal.
FT DOMAIN 2456 2751 Nonhelical region (NC3).
FT DOMAIN 2752 2899 Triple-helical region (COL2) with 1
FT DOMAIN 2900 2942 Triple-helical region (NC2).
FT DOMAIN 2943 3045 Triple-helical region (COL1) with 2
FT DOMAIN 3046 3119 Nonhelical region (NC1).
FT SITE 866 868 Cell attachment site (Potential).
FT SITE 2784 2786 Cell attachment site (Potential).
FT SITE 2896 2898 Cell attachment site (Potential).
FT MOD_RES 2945 2945 Hydroxyproline (By similarity).
FT MOD_RES 2948 2948 Hydroxyproline (By similarity).
FT MOD_RES 2951 2951 Hydroxyproline (By similarity).
FT MOD_RES 2960 2960 Hydroxyproline (By similarity).
FT MOD_RES 2966 2966 Hydroxyproline (By similarity).
FT MOD_RES 2969 2969 Hydroxyproline (By similarity).
FT MOD_RES 2972 2972 Hydroxyproline (By similarity).
FT MOD_RES 2984 2984 Hydroxyproline (By similarity).
FT MOD_RES 3001 3001 Hydroxyproline (By similarity).
FT MOD_RES 3004 3004 Hydroxyproline (By similarity).
FT MOD_RES 3015 3015 Hydroxyproline (By similarity).
FT MOD_RES 3024 3024 Hydroxyproline (By similarity).
FT MOD_RES 3027 3027 Hydroxyproline (By similarity).
FT MOD_RES 3030 3030 Hydroxyproline (By similarity).
FT DOMAIN 869 872 Poly-Thr.
FT CARBOHYD 704 704 O-linked (GlcNAc-. ) (Potential).
FT CARBOHYD 802 802 O-linked (Xyl-. ) (Chondroitin sulfate)
FT CARBOHYD 893 893 O-linked (Xyl-. ) (Chondroitin sulfate)
FT CARBOHYD 985 985 O-linked (Xyl-. ) (Chondroitin sulfate)
FT CARBOHYD 1769 1769 N-linked (GlcNAc-. ) (Potential).
FT CARBOHYD 2212 2212 N-linked (GlcNAc-. ) (Potential).
FT CARBOHYD 2533 2533 N-linked (GlcNAc-. ) (Potential).
FT CARBOHYD 2684 2684 N-linked (GlcNAc-. ) (Potential).
FT VARSPLIC 25 1190 Missing (in isoform XIIB-1 and isoform
FT XIIB-2).
FT VARSPLIC 3062 3064 /FTId=VSP_001150.
FT VARSPLIC 3065 3119 Epy -> GSG (in isoform XIIB-2 and isoform
FT XIIB-2).
FT VARSPLIC 3065 3119 /FTId=VSP_001151.
FT VARSPLIC 3065 3119 Missing (in isoform XIIB-2 and isoform
FT XIIB-2).
SQ SEQUENCE 3119 AA, 340239 MW, 981F999C86AB3251 CRC64,
Query Match 14.3%; Score 139; DB 1; Length 3119;
Best Local Similarity 30.3%; Pred. No. 0.015;
Matches 60; Conservative 37; Mismatches 75; Indels 26; Gaps 12;
```

```
AC Q8T5C2;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Proximal thread matrix protein 1.
OS Mytilus galloprovincialis (Mediterranean mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Periomorphia; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_TaxID=29158;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2231036; PubMed=12425661; DOI=10.1021/bm0255903;
RA Sun C., Lucas J.M., Waite J.H.;
RT "Collagen-binding matrix proteins from elastomeric extraorganic
RT byssal fibers.";
RL Biomacromolecules 3:1240-1248 (2002).
DR EMBL; AY053391; AL117974.1; -.
DR HSSP; P20701; IMUN.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF00092; VMA; 2.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00327; VMA; 2.
DR PROSITE; PS50234; WMA; 2.
KW Matrix protein.
SQ SEQUENCE 453 AA; 48784 MW; D60497F5C0C51BED CRC64;
Query Match 14.2%; Score 138; DB 2; Length 453;
Best Local Similarity 25.3%; Pred. No. 0.0019;
Matches 50; Conservative 44; Mismatches 78; Indels 26; Gaps 12;
```


CC -1- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-M
 CC associates with beta-2.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: Predominantly expressed in monocytes and
 CC granulocytes.
 CC -1- DOMAIN: The integrin I-domain (insert) is a WFPA domain. Integrins
 CC with I-domains do not undergo protease cleavage.
 CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
 CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -1- SIMILARITY: Contains 1 FG-GAP domain.
 CC -1- DATABASE: NAME=PROV; NOTE=CD guide CD11b entry;
 CC WWW=http://www.ncbi.nlm.nih.gov/prov/cd/cd11b.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.emb-rib.ch/announce/
 CC or send an email to license@emb-rib.ch).
 CC -----
 DR EMBL/ J03925; AAA59544.1; -;
 DR EMBL/ M18044; AAA59491.1; -;
 DR EMBL/ J04145; AAA59903.1; -;
 DR EMBL/ S52227; AAB24821.1; -;
 DR EMBL/ S52152; AAB24821.1; JOINED.
 DR EMBL/ S52153; AAB24821.1; JOINED.
 DR EMBL/ S52154; AAB24821.1; JOINED.
 DR EMBL/ S52155; AAB24821.1; JOINED.
 DR EMBL/ S52157; AAB24821.1; JOINED.
 DR EMBL/ S52159; AAB24821.1; JOINED.
 DR EMBL/ S52161; AAB24821.1; JOINED.
 DR EMBL/ S52164; AAB24821.1; JOINED.
 DR EMBL/ S52165; AAB24821.1; JOINED.
 DR EMBL/ S52167; AAB24821.1; JOINED.
 DR EMBL/ S52170; AAB24821.1; JOINED.
 DR EMBL/ S52173; AAB24821.1; JOINED.
 DR EMBL/ S52180; AAB24821.1; JOINED.
 DR EMBL/ S52181; AAB24821.1; JOINED.
 DR EMBL/ S52184; AAB24821.1; JOINED.
 DR EMBL/ S52189; AAB24821.1; JOINED.
 DR EMBL/ S52191; AAB24821.1; JOINED.
 DR EMBL/ S52192; AAB24821.1; JOINED.
 DR EMBL/ S52203; AAB24821.1; JOINED.
 DR EMBL/ S52212; AAB24821.1; JOINED.
 DR EMBL/ S52216; AAB24821.1; JOINED.
 DR EMBL/ S52219; AAB24821.1; JOINED.
 DR EMBL/ S52220; AAB24821.1; JOINED.
 DR EMBL/ S52221; AAB24821.1; JOINED.
 DR EMBL/ S52226; AAB24821.1; JOINED.
 DR EMBL/ M76724; AAA58410.1; -;
 DR EMBL/ M84477; AAA51960.1; -;
 DR PIR/ A31108; RWHUB.
 DR PDB/ 1A8X; Model; @=17-1152.
 DR PDB/ 1BHO; X-ray; 1/2=-.
 DR PDB/ 1BDQ; X-ray; 1/2=-.
 DR PDB/ 1IDN; X-ray; 1/2=-.
 DR PDB/ 1IDO; X-ray; @=140-331.
 DR PDB/ 1JLM; X-ray; @=143-334.
 DR PDB/ 1MLU; X-ray; A=137-331.
 DR PDB/ 1MP7; X-ray; A=144-337.
 DR PDB/ 1N9Z; X-ray; A=140-335.
 DR PDB/ 1N9Z; X-ray; A=144-345.
 DR PDB/ 1N9Z; X-ray; A=144-345.
 DR Genew/ HGNC:6149; ITGAM.
 DR MIM/ 120980; -;
 DR GO/ GO:0008305; C:integrin complex; TAS.
 DR GO/ GO:0007155; P:cell adhesion; TAS.
 DR Interpro/ IPR000413; Integrin_alpha.
 DR Interpro/ IPR002035; VWF_A.

DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00357; Integrin_alpha; 1.
 DR Pfam; PF00092; VWA; 1.
 DR PRINTS; PRO1185; INTEGRINA.
 DR PRINTS; PRO0453; VWFADOMAIN.
 DR SMART; SM00191; Int_alpha; 5.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS0234; VWFPA; 1.
 DR 3D-structure; Calcium; Cell adhesion; Direct protein sequencing;
 KW Glycoprotein; Integrin; Magnesium; Receptor; Repeat; Signal;
 KW Transmembrane.
 FT SIGNAL 1 16
 FT CHAIN 17 1152 Integrin alpha-M.
 Query Match 13.7%; Score 132.5; DB 1; Length 1152;
 Best Local Similarity 26.3%; Pred. No. 0.017;
 Matches 52; Conservative 41; Mismatches 66; Indels 39; Gaps 11;
 QY 4 DLYFLDKSGSVL-HHNNELIYFVEQLAHKFIQPLMSFTVSTGRTTLMKLTED--- 58
 Db 150 DIAFLIDSGSILPHDRRMKEFVST-----VMEQLKSKTLEFS-----IMQYSEPRRIH 199
 QY 59 -----RQIQGLEELQKVPBGSDTYHGEFRASQIYYENRCQYRTA-SVITALT 109
 Db 200 FTFKEFQNNPNPSLVPITQL--GRTHATGIRKVRLEFNITGARKNAFKILVIT 257
 QY 110 DGEIHEDLPFYSE--REANRSRD/GAIVYCGVVDFFNETOLAR-----IADS--KDHVFP 160
 Db 258 DGEKFGDPLGVEDYIPADNR--GVIRYIVGVADARSRSEKROELNTASKPRDHFQ 314
 QY 161 VNDGFOALGGIHSILKK 178
 Db 315 VNN-FEALKTIONQLREK 331
 RESULT 14
 CAIC HUMAN STANDARD; PRT; 3063 AA.
 AC 099715; 099716;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Collagen alpha 1(XII) chain precursor.
 GN Name=COL12A1;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RP [1]
 RX MEDLINE=97288521; PubMed=9143499; DOI=10.1006/geno.1997.4638;
 RA Gerecke D.R., Olson P.F., Koch M., Knoll J.H.M., Taylor R.
 RA Hudson D.L., Champilaud M.-F., Olsen B.R., Bursen R.E.;
 RT "Complete primary structure of two splice variants of collagen XII,
 RT and assignment of alpha 1(XII) collagen (COL12A1), alpha 1(IX)
 RT collagen (COL9A1), and alpha 1(XIX) collagen (COL19A1) to human
 RT chromosome 6q12-q13."
 RT Genomics 41:236-242(1997).
 RL Genomics 41:236-242(1997).
 CC -1- FUNCTION: Type XII collagen interacts with type I collagen-
 CC containing fibrils. The COL1 domain could be associated with the
 CC surface of the fibrils, and the COL2 and NC3 domains may be
 CC localized in the pericollagen matrix (By similarity).
 CC -1- SUBUNIT: Trimer of identical chains each containing 190 kDa of
 CC nontriple-helical sequences.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Comment=The final tissue form of collagen XII may contain
 CC homotrimers of either isoform long or isoform short or any
 CC combination of isoform long and isoform short;
 CC Name=Long;
 CC IsoId=099715-1; Sequence=Displayed;
 CC Name=Short;

CC	Isoid-099715-2; Sequence-VSP_001149;	
CC	-1- TISSUE SPECIFICITY: Found in collagen I-containing tissues: both	
CC	short and long isoforms appear in amnion, chorion, skeletal	
CC	muscle, small intestine, and in cell culture of dermal	
CC	fibroblasts, keratinocytes and endothelial cells. Only the short	
CC	isoform is found in lung, placenta, kidney and a squamous cell	
CC	carcinoma cell line.	
CC	-1- Ptm: The triple-helical tail is stabilized by disulfide bonds at	
CC	each end (By similarity).	
CC	-1- Ptm: Prolines at the third position of the tripeptide repeating	
CC	unit (G-X-Y) are hydroxylated in some or all of the chains (By	
CC	similarity).	
CC	-1- Ptm: O-glycosylation of isoform long; glycosaminoglycan of	
CC	chondroitin-sulfate type (By similarity).	
CC	-1- SIMILARITY: Belongs to the fibril-associated collagens with	
CC	interrupted helices (FACIT) family.	
CC	-1- SIMILARITY: Contains 18 fibronectin type III domains.	
CC	-1- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.	
CC	-1- SIMILARITY: Contains 4 VWFA domains.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/	
CC	or send an email to license@isb-sib.ch).	
CC	-----	
DR	EMBL; U73778; AAC51244.1; -	
DR	EMBL; U73779; AAD40463.1; -	
DR	HSSP; P18614; IMHP.	
DR	Genew; HGNC:2188; COL12A1.	
DR	MIM; 120320; -	
DR	GO; GO:0005595; C:collagen type XII; TAS.	
DR	GO; GO:0001501; P:skeletal development; TAS.	
DR	InterPro; IPR008160; Collagen.	
DR	InterPro; IPR008985; Coma_Like_Jec_g1.	
DR	InterPro; IPR003961; FN_III.	
DR	InterPro; IPR008957; FN_III-1like.	
DR	InterPro; IPR002035; VWF_A.	
DR	Pfam; PF01391; Collagen; 4.	
DR	Pfam; PF00041; FN3_18.	
DR	Pfam; PF02210; TSP_N; 1.	
DR	Pfam; PF00092; VWF; 4.	
DR	PRINTS; PR00453; VWFADOMAIN.	
DR	PROSITE; PS00853; FN3_18.	
DR	PROSITE; PS50234; VWFA; 4.	
KW	Alternative splicing; Cell adhesion; Collagen;	
KW	Direct protein sequencing; Extracellular matrix; Glycoprotein;	
KW	Hydroxylation; Repeat; Signal; Structural protein.	
KW	Potential.	
FT	CHAIN	1 24
FT	COLLAGEN ALPHA 1(XII) CHAIN.	
FT	25 3063	
FT	140 112	
FT	DOMAIN	140 316
FT	333 422	
FT	DOMAIN	440 616
FT	631 719	
FT	DOMAIN	722 810
FT	813 901	
FT	DOMAIN	904 993
FT	995 1083	
FT	DOMAIN	1086 1175
FT	1199 1371	
FT	DOMAIN	1384 1472
FT	1474 1563	
FT	DOMAIN	1565 1652
FT	1654 1743	
FT	DOMAIN	1752 1841
FT	1843 1931	
FT	DOMAIN	1933 2022
FT	2024 2113	
FT	DOMAIN	2115 2202
FT	2206 2290	
FT	DOMAIN	2206 2290

FT	DOMAIN	2223	2496	VFPA 4.
FT	DOMAIN	2520	2712	TSP N-terminal.
FT	DOMAIN	2451	2746	Nonhelical region (NC3).
FT	DOMAIN	2747	2898	Triple-helical region (COL2) with 1 imperfection.
FT	DOMAIN	2899	2941	Nonhelical region (NC2).
FT	DOMAIN	2942	3044	Triple-helical region (COL1) with 2 imperfections.
FT	DOMAIN	3045	3063	Nonhelical region (NC1).
FT	SITE	862	864	Cell attachment site (Potential).
FT	SITE	2779	2781	Cell attachment site (Potential).
FT	SITE	2895	2897	Cell attachment site (Potential).
FT	MOD_RES	2944	2944	Hydroxyproline (By similarity).
FT	MOD_RES	2947	2947	Hydroxyproline (By similarity).
FT	MOD_RES	2950	2950	Hydroxyproline (By similarity).
FT	MOD_RES	2959	2959	Hydroxyproline (By similarity).
FT	MOD_RES	2965	2965	Hydroxyproline (By similarity).
FT	MOD_RES	2968	2968	Hydroxyproline (By similarity).
FT	MOD_RES	2971	2971	Hydroxyproline (By similarity).
FT	MOD_RES	2983	2983	Hydroxyproline (By similarity).
FT	MOD_RES	3000	3000	Hydroxyproline (By similarity).
FT	MOD_RES	3003	3003	Hydroxyproline (By similarity).
FT	MOD_RES	3014	3014	Hydroxyproline (By similarity).
FT	MOD_RES	3023	3023	Hydroxyproline (By similarity).
FT	MOD_RES	3026	3026	Hydroxyproline (By similarity).
FT	MOD_RES	3029	3029	Hydroxyproline (By similarity).
FT	CARBOHYD	700	700	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	798	798	O-linked (Xyl. . .) (chondroitin sulfate) (Potential).
FT	CARBOHYD	889	889	O-linked (Xyl. . .) (chondroitin sulfate) (Potential).
FT	CARBOHYD	981	981	O-linked (Xyl. . .) (chondroitin sulfate) (Potential).
FT	CARBOHYD	1763	1763	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	2206	2206	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	2528	2528	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	2679	2679	N-linked (GlcNAc. . .) (Potential).
FT	VARSPLIC	25	1188	Missing (in isoform Short).
SEQ	SEQUENCE	3063 AA; 333189 MW; 75FEA76FA8B48293 CRC64;		
Query Match	Similarity	13.5%;	Score 131, DB 1; Length 3063;	
Best Local	Similarity	29.8%;	Pred. No. 0.07;	
Matches	Conservative	38;	Mismatches 75; Indels 26; Gaps 13	
Qy	4	DLYETILKSGSV-LHHNNEIYYPFEOIAHKR-ISP-QLRMSFIYVST---RGTTIMKLTJE	57	
Db	440	DIVLVLGVSSISIGIANFKVNAFLEVLVKSREISDNRVQISLVQYSRDPHFETLKKTTK	499	
Qy	58	DREOIRGLSELTAKVLP--GGDTVMHSGFERASEIQIYENKQGYRT--ASYVIALTDGEL	113	
Db	500	VEDI-----EAINTFPYRGSTTNGKAMTYVRREKIPVPSK-GGSNNPKMILLTDGK	552	
Qy	114	HEDLVFYSEREANSRDGLAIVCVGVKDFNETOLARIAD--SKDHFVPNDGFOALGI	171	
Db	553	SSDAF---RDPFAIKLRNSDVEIFAVGVDAVRSELEAIASPPAEHVFTEVD-FDAFQRI	608	
Qy	172	IHSILKKSCTEI---LAA	186	
Db	609	SFEITQISICRIEQELAA	626	
RESULT 15				
Q8C6K9				
AC	Q8C6K9	PRELIMINARY;	PRT; 1182 AA.	
DT	01-MAR-2003 (TREMBLrel. 23, Created)			
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)			
DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)			
DE	Mus musculus 2 days pregnant adult female ovary cDNA, RIKEN full-length enriched library, clone:E330019B14 product:hypothetical von Willebrand factor type A domain containing protein, full insert sequence.			

GN Name=E330026802R1k;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning";
RL Mech. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RA The FANTOM Consortium.
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs.";
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RL prepare full-length cDNA libraries for rapid discovery of new genes.";
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagasaka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishii K., Kitsuai T., Tashiro H., Itoh M.,
RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsunoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara S., Watanishi M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RL sequencing pipeline with 384 multicapillary sequencer.";
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX Adegchi J., Aizawa K., Akimura T., Arahawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiroaka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kaubawa T.,
RA Kachi H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
RA Kurihara C., Maebayama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ono M., Ohsato N., Okazaki Y.,
RA Saeki R., Satoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Saeki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toyu T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/Genbank/DBD databases.
DR EMBL; AK054356; BAC35749.1; -;
DR HSP; P11215; IMP7.
DR MGD; MGI:2444259; E330026802R1k.
DR InterPro; IPRO02035; VWF_A.
DR Pfam; PF00092; VMA_6.
DR PRINTS; PR00493; VWFADOMAIN.
DR SMART; SM00327; VMA_6.
DR PROSITE; PSS0234; VMPA_6.

[illegible]

Search completed: June 13, 2005, 20:03:15
Job time : 63.775 secs

GN Name=E330026802R1k;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning";
RL Mech. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RA The FANTOM Consortium.
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs.";
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RL prepare full-length cDNA libraries for rapid discovery of new genes.";
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagasaka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishii K., Kitsuai T., Tashiro H., Itoh M.,
RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsunoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara S., Watanabe M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RL sequencing pipeline with 384 multicapillary sequencer.";
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX Adegchi J., Aizawa K., Akimura T., Arahawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiroaka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kaubawa T.,
RA Kachi H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
RA Kurihara C., Maebayama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ono M., Ohsato N., Okazaki Y.,
RA Saeki R., Satoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Saeki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toyu T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBD databases.
DR EMBL; AK054356; BAC35749.1; -;
DR HSP; P11215; IMP7.
DR MGD; MGI:2444259; E330026802R1k.
DR InterPro; IPRO02035; VWF_A.
DR Pfam; PF00092; VMA_6.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00327; VMA_6.
DR PROSITE; PSS0234; VWFPA_6.

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 19:37:17 ; Search time 67.0133 Seconds
(without alignments)
1044.623 Million cell updates/sec

Title: US-09-970-076-2_COPY_42_222
Perfect score: 943
Sequence: 1 GFDLYRILDKSGSVLHWMR.....DGFQALQGITHSLKSCIR 181

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:.*
1: geneseqp19808:.*
2: geneseqp19908:.*
3: geneseqp20008:.*
4: geneseqp20018:.*
5: geneseqp20028:.*
6: geneseqp20038:.*
7: geneseqp20038:.*
8: geneseqp20048:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	943	100.0	297	4 AAM38976	Aam38976 Human pol
2	943	100.0	328	7 ADI00558	Adi00558 Human TAN
3	943	100.0	328	7 ADM64584	Adm64584 Human TAN
4	943	100.0	333	3 AAB01422	Aab01422 Human TAN
5	943	100.0	333	5 ABP54905	Abp54905 Human ant
6	943	100.0	333	7 ADI00534	Adi00534 Human TAN
7	943	100.0	333	7 ADM64568	Adm64568 Human von
8	943	100.0	342	7 ADI00554	Adi00554 Human TAN
9	943	100.0	342	7 ADM64580	Adm64580 Human TAN
10	943	100.0	345	7 ADI00556	Adi00556 Human TAN
11	943	100.0	345	7 ADM64582	Adm64582 Human TAN
12	943	100.0	368	5 ABP54903	Abp54903 Human ant
13	943	100.0	384	7 ADM64586	Adm64586 TANGO197
14	943	100.0	403	4 AAB01439	Aae01439 Human gen
15	943	100.0	403	5 ABG63874	Abg63874 Human alb
16	943	100.0	403	8 ADL77139	Adl77139 Albumin f
17	943	100.0	460	7 ADI00560	Adi00560 Human TAN
18	943	100.0	479	7 ADI00564	Adi00564 Human TAN
19	943	100.0	504	7 ADI00566	Adi00566 Human TAN
20	943	100.0	529	7 ADI00568	Adi00568 Human TAN
21	943	100.0	540	7 ADI00544	Adi00544 Human TAN
22	943	100.0	540	7 ADM64588	Adm64588 TANGO197
23	943	100.0	549	7 ADI00546	Adi00546 Human TAN
24	943	100.0	549	7 ADI00542	Adi00542 Human TAN
25	943	100.0	549	7 ADI00548	Adi00548 Human TAN

26	943	100.0	549	7 ADM64590	Adm64590 TANGO197
27	943	100.0	549	7 ADM64592	Adm64592 TANGO197
28	943	100.0	551	7 ADI00550	Adi00550 Human TAN
29	943	100.0	551	7 ADM64576	Adm64576 Mouse TAN
30	943	100.0	564	5 ABB90750	Abb90750 Human Tum
31	943	100.0	564	5 ABB90724	Abb90724 Human Tum
32	943	100.0	564	5 ABP54904	Abp54904 Human ant
33	943	100.0	564	5 ABP54457	Abp54457 Human Tum
34	943	100.0	564	6 ABUS4431	Abu54431 Human Tum
35	943	100.0	564	7 ADI00552	Adi00552 Human TAN
36	943	100.0	564	7 ADJ70017	Adj70017 Human hea
37	943	100.0	564	7 ADM64578	Adm64578 Human TAN
38	943	100.0	564	8 ADR48216	Adr48216 Human Tum
39	938	99.5	403	4 AAB01469	Aae01469 Human gen
40	938	99.5	403	5 ABG63873	Abg63873 Human alb
41	938	99.5	403	8 ADL77138	Adl77138 Albumin f
42	934	99.0	562	5 ABB90731	Abb90731 Mouse Tum
43	934	99.0	562	5 ABB90785	Abb90785 Mouse Tum
44	934	99.0	562	6 ABUS4492	Abu54492 Mouse Tum
45	934	99.0	562	6 ABUS4438	Abu54438 Mouse Tum

ALIGNMENTS

RESULT 1
AAM38976
ID AAM38976 standard; protein; 297 AA.
XX
AC AAM38976;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 2121.
XX
KW Human, noctropic; immunosuppressant; cyostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN WO20015312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WC-US034263.
XX
PR 23-DEC-1999; 99US-00471275.
XX
PR 21-JAN-2000; 2000US-00488725.
XX
PR 25-APR-2000; 2000US-00552317.
XX
PR 20-JUN-2000; 2000US-00598042.
XX
PR 19-JUL-2000; 2000US-00620312.
XX
PR 03-AUG-2000; 2000US-00653450.
XX
PR 14-SEP-2000; 2000US-00662191.
XX
PR 19-OCT-2000; 2000US-00693036.
XX
PR 29-NOV-2000; 2000US-00727344.
XX
XX (HYSE-) HYSBO INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
XX Zhou P, Goodrich R, Drmanac RT;
XX WPI; 2001-442253/47.
XX N-PSDB; AA158132.
XX Novel nucleic acids and polypeptides, useful for treating disorders such
XX as central nervous system injuries.
XX Example 4; SEQ ID NO 2121; 10078bp; English.
PS

XX The invention relates to human nucleic acids (AA157798-AA161369) and the
CC encoded polypeptides (AA038642-AA042213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification

XX
SQ Sequence 297 AA;

Query Match 100.0%; Score 943; DB 4; Length 297;
Best Local Similarity 100.0%; Pred. No. 1.7e-96;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFDLYFLDKSGSVLHHMNEIYFVEQLAHKFTSPQLRMSFIVSTGTTLMLKLTREQ 60
DB 42 GFDLYFLDKSGSVLHHMNEIYFVEQLAHKFTSPQLRMSFIVSTGTTLMLKLTREQ 101

QY 61 IROGLELQVLPQGDVTYMHGFEFRASEQIYENRQGYRTASVIALTDGELHEDLFFYS 120
DB 102 IROGLELQVLPQGDVTYMHGFEFRASEQIYENRQGYRTASVIALTDGELHEDLFFYS 161

QY 121 ERENANRSDIGAIVYCVKDFNETQARIADSKDHFVPVNDGFQALOGIHSILKKSCT 180
DB 162 ERENANRSDIGAIVYCVKDFNETQARIADSKDHFVPVNDGFQALOGIHSILKKSCT 221

QY 181 E 181
DB 222 E 222

RESULT 2
ID ADI00558
ADI00558 standard; protein; 328 AA.

AC ADI00558;
XX
DT 22-APR-2004 (first entry)
XX
DE Human TANGO 197 HisTag fusion protein - plasmid p0615.
XX
KW fusion; von Willebrand factor A-like domain; vWF; antibacterial;
KW cutaneous; inhalation anthrax; human; TANGO 197 HisTag fusion; mutant;
XX plasmid p0615; mutcin.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN US2003144193-A1.
XX
PD 31-JUL-2003.
XX
PF 24-JUL-2002; 2002US-00201292.
XX
PR 20-DEC-2001; 2001US-00038307.
XX
PA (ROTT/) ROTTMAN J B.
PA (OKEE/) O'KEEFE T L.
PA (OZKA/) OZKAYNAK E.
PA (HEAL/) HEALEY J J.
XX
PI Rottman JB, O'Keefe TL, Ozkaynak E, Healey JJ;
XX
DR WPI, 2003-720708/68.

DR N-PSDB; ADI00557.
XX
XX New TANGO 197 and/or TANGO 216 fusion polypeptides for preventing or
PT ameliorating symptoms of anthrax comprises a von Willebrand factor A-like
PT domain (vWF) amino acid sequence and an amino acid sequence heterologous
PT to the vWF.
XX
PS Claim 45; SEQ ID NO 26; 86pp; English.
XX
XX The invention relates to a novel fusion polypeptide comprising a von
CC Willebrand factor A-like domain (vWF) amino acid sequence and an amino
CC acid sequence heterologous to the vWF. The polypeptide of the invention
CC demonstrates antibacterial activities whilst the composition and method
CC may be useful in preventing or ameliorating the symptoms of cutaneous
CC and/or inhalation anthrax. The current sequence is that of the human
CC TANGO 197 HisTag fusion protein of the invention.

XX
SQ Sequence 328 AA;

Query Match 100.0%; Score 943; DB 7; Length 328;
Best Local Similarity 100.0%; Pred. No. 2e-96;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFDLYFLDKSGSVLHHMNEIYFVEQLAHKFTSPQLRMSFIVSTGTTLMLKLTREQ 60
DB 42 GFDLYFLDKSGSVLHHMNEIYFVEQLAHKFTSPQLRMSFIVSTGTTLMLKLTREQ 101

QY 61 IROGLELQVLPQGDVTYMHGFEFRASEQIYENRQGYRTASVIALTDGELHEDLFFYS 120
DB 102 IROGLELQVLPQGDVTYMHGFEFRASEQIYENRQGYRTASVIALTDGELHEDLFFYS 161

QY 121 ERENANRSDIGAIVYCVKDFNETQARIADSKDHFVPVNDGFQALOGIHSILKKSCT 180
DB 162 ERENANRSDIGAIVYCVKDFNETQARIADSKDHFVPVNDGFQALOGIHSILKKSCT 221

QY 181 E 181
DB 222 E 222

RESULT 3
ID ADM64584
ADM64584 standard; protein; 328 AA.

AC ADM64584;
XX
DT 03-JUN-2004 (first entry)
XX
DE Human TANGO197-His tag fusion protein #2.
XX
XX antibacterial; gene therapy;
KW von Willebrand factor A-like domain amino acid sequence;
KW vWF amino acid sequence; anthrax; Bacillus anthracis; cutaneous anthrax;
XX inhalation anthrax; human; TANGO197; his tag; fusion protein.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN US2003134786-A1.
XX
PD 17-JUL-2003.
XX
PF 20-DEC-2001; 2001US-00038307.
XX
PR 20-DEC-2001; 2001US-00038307.
XX
PA (ROTT/) ROTTMAN J B.
PA (OKEE/) O'KEEFE T L.
PA (OZKA/) OZKAYNAK E.
PA (HEAL/) HEALEY J J.
XX
PI Rottman JB, O'Keefe TL, Ozkaynak E, Healey JJ;
XX
XX

DR WPI, 2003-829643/77.
DR N-PSDB; ADM64583.
XX
XX
PT New TANGO 197 and/or TANGO 216 fusion polypeptides for preventing or
PT ameliorating symptoms of anthrax comprises a von Willebrand factor A-like
PT domain (vWF) amino acid sequence and an amino acid sequence heterologous
PT to the vWF.
XX
XX
PS Claim 44; SEQ ID NO 26; 64bp; English.
XX
CC The invention describes a fusion polypeptide comprising a von Willebrand
CC factor A-like domain (vWF) amino acid sequence and an amino acid sequence
CC heterologous to the vWF. Also described are: a method of preventing or
CC ameliorating a symptom of anthrax in a subject thought to be at risk for
CC exposure to or suspected of having been exposed to Bacillus anthracis;
CC and a pharmaceutical composition comprising the novel fusion polypeptide.
CC The composition and method are useful in preventing or ameliorating
CC symptoms of cutaneous and/or inhalation anthrax. This is the amino acid
CC sequence of a fusion protein comprising mature human TANGO197, thrombin
CC cleavage site and his tag that can be used to treat exposure to or
CC prevent a symptom of anthrax.
XX
SQ Sequence 328 AA;
XX
Query Match 100.0%; Score 943; DB 7; Length 328;
Best Local Similarity 100.0%; Pred. No. 2e-96;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 GFDLYFLDKSGSVLHMNNEIYFVEQLAKHFI SPQLRMSFIYVSTGTTLMKLTEDREQ 60
DB 42 GFDLYFLDKSGSVLHMNNEIYFVEQLAKHFI SPQLRMSFIYVSTGTTLMKLTEDREQ 101
XX
QY 61 IROGLEELQKVLPGGDTYMHGEPFRASEQIYENRQGRTRASVIALTDGELHEDLFFYS 120
DB 102 IROGLEELQKVLPGGDTYMHGEPFRASEQIYENRQGRTRASVIALTDGELHEDLFFYS 161
XX
QY 121 EREANRSRDIGAIYVCVGDVFNFTOLARIADSKDHVPVNDGFOALQGIHSLTKKSCI 180
DB 162 EREANRSRDIGAIYVCVGDVFNFTOLARIADSKDHVPVNDGFOALQGIHSLTKKSCI 221
XX
QY 181 E 181
DB 222 E 222
XX
RESULT 4
AAB01422
ID AAB01422 standard; protein; 333 AA.
XX
AC AAB01422;
XX
DT 20-OCT-2000 (first entry)
XX
DE Human TANGO 197.
XX
XX TANGO, 128; 140; 197; 212; 224; 239; modulating agent; asthma;
XX graft versus-host diseases; rheumatoid arthritis; psoriasis;
XX inflammatory bowel diseases; septic shock; ulcerative colitis;
XX Crohn's disease; chronic myelogenous leukemia; cancer; liver disease;
XX Hodgkin's disease; osteoarthritis; Lyme's disease; cachexia;
XX autoimmune disease; myasthenia gravis; autoimmune diabetes;
XX systemic lupus erythematosus; transgenic animal; diagnosis; prognosis;
XX prophylactic; therapeutic; human.
XX
OS Homo sapiens.
XX
PN MO200039284-A1.
XX
PD 06-JUL-2000.
XX
PF 23-DEC-1999; 99MO-US031025.
XX
PR 30-DEC-1998; 98US-00223546.

XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Holtzman DA;
XX
XX WPI, 2000-465743/40.
DR N-PSDB; AAA47455.
XX
XX
PT Novel nucleic acid sequences encoding TANGO-128, 140, 197, 212, 213, 224
PT and 239 polypeptides useful for the treatment of asthma, rheumatoid
PT arthritis, psoriasis and autoimmune diseases.
XX
XX
PS Claim 8; Fig 4; 209bp; English.
XX
XX Nucleic acids encoding TANGO polypeptides are useful as modulating agents
CC for regulating cellular processes like asthma, graft versus-host
CC diseases, rheumatoid arthritis, psoriasis, inflammatory bowel disease,
CC septic shock, ulcerative colitis, Crohn's disease, chronic myelogenous
CC leukemia, cancer, liver disease, Hodgkin's disease, osteoarthritis,
CC Lyme's disease, cachexia and autoimmune diseases e.g. myasthenia gravis,
CC autoimmune diabetes and systemic lupus erythematosus. The nucleic acids
CC are also useful for producing transgenic animals and the TANGO
CC polypeptides themselves. Partial TANGO-128, 140, 197, 212, 213, 224, 239
CC sequences are useful in forensic biology, for diagnostic assays,
CC prognostic assays, pharmacogenomics and for monitoring clinical trials.
CC TANGO polypeptides are suitable for both prophylactic and therapeutic
CC methods for treating a subject at risk of a disorder or having a disorder
CC associated with aberrant TANGO expression. A wide range of cellular
CC disorders can be treated
XX
SQ Sequence 333 AA;
XX
Query Match 100.0%; Score 943; DB 3; Length 333;
Best Local Similarity 100.0%; Pred. No. 2e-96;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 GFDLYFLDKSGSVLHMNNEIYFVEQLAKHFI SPQLRMSFIYVSTGTTLMKLTEDREQ 60
DB 42 GFDLYFLDKSGSVLHMNNEIYFVEQLAKHFI SPQLRMSFIYVSTGTTLMKLTEDREQ 101
XX
QY 61 IROGLEELQKVLPGGDTYMHGEPFRASEQIYENRQGRTRASVIALTDGELHEDLFFYS 120
DB 102 IROGLEELQKVLPGGDTYMHGEPFRASEQIYENRQGRTRASVIALTDGELHEDLFFYS 161
XX
QY 121 EREANRSRDIGAIYVCVGDVFNFTOLARIADSKDHVPVNDGFOALQGIHSLTKKSCI 180
DB 162 EREANRSRDIGAIYVCVGDVFNFTOLARIADSKDHVPVNDGFOALQGIHSLTKKSCI 221
XX
QY 181 E 181
DB 222 E 222
XX
RESULT 5
ABP54905
ID ABP54905 standard; protein; 333 AA.
XX
AC ABP54905;
XX
DT 08-JAN-2003 (first entry)
XX
DE Human anthrax toxin receptor.
XX
XX Anthrax; toxin; receptor; human; antibacterial.
XX
OS Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX FT Peptide 1..27
XX FT Protein /label= Signal_peptide
XX FT 28..333
XX FT /label= Mature_protein
XX

CC Anthrax toxin receptor (ATR), as predicted from an isolated cDNA clone.
CC Anthrax toxin protective antigen (PA) binds to the ATR at a von
CC Willebrand factor A domain located in the extracellular domain of ATR.
CC The invention provides ATR polypeptides and polynucleotides, vectors,
CC host cells, and transgenic and knock-out animals. It also provides
CC methods for identifying molecules that bind the ATR and which reduce the
CC toxicity of anthrax toxin. A claimed method for treating anthrax in a
CC human or animal involves administering an agent that inhibits binding
CC between PA and ATR at a level effective to reduce the severity of
CC anthrax. Suitable agents include ATR or a PA-binding fragment of ATR, a
CC PA-binding polypeptide at least 80% identical to these, a fusion protein,
CC a monoclonal or polyclonal antibody, a polysaccharide, a lipid or a
CC nucleic acid
CC
SQ Sequence 368 AA;

Query Match 100.0%; Score 943; DB 5; Length 368;
Best Local Similarity 100.0%; Pred. No. 2.3e-96;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFDLYFLDKSGSVLHHMNEIYFVEQLAHKFI SPQLRMSFIVSTRGTLMLKLTEDREQ 60
DB 42 GFDLYFLDKSGSVLHHMNEIYFVEQLAHKFI SPQLRMSFIVSTRGTLMLKLTEDREQ 101
QY 61 IRGGLSELQKVLPGSDTYMHGFERASEQIYYENRQGYRTASVIALTDGELHEDLFFYS 120
DB 102 IRGGLSELQKVLPGSDTYMHGFERASEQIYYENRQGYRTASVIALTDGELHEDLFFYS 161
QY 121 EREANRSRDLGALVYCVGVDFNETQLARIADSKDHVPFVNDGFQALQGIHSLKKSCT 180
DB 162 EREANRSRDLGALVYCVGVDFNETQLARIADSKDHVPFVNDGFQALQGIHSLKKSCT 221
QY 181 E 181
DB 222 E 222

RESULT 13
ADM64586
ID ADM64586 standard; protein; 384 AA.
XX
AC ADM64586;
XX
DT 03-JUN-2004 (first entry)
XX
DE TANGO197 extracellular domain-mutant IgG Fc fusion protein #1.
XX
XX antibacterial; gene therapy;
XX von Willebrand factor A-like domain amino acid sequence;
XX WVF amino acid sequence; anthrax; Bacillus anthracis; cutaneous anthrax;
XX Inhalation anthrax; human; TANGO197; immunoglobulin G; IgG;
XX fragment of crystallisation; Fc.
XX
XX Homo sapiens.
OS Synthetic.
XX
XX US2003134786-A1.
XX
XX 17-JUL-2003.
XX
XX 20-DEC-2001; 2001US-00038307.
XX
XX 20-DEC-2001; 2001US-00038307.
XX
XX 20-DEC-2001; 2001US-00038307.
XX
XX (ROT/) ROTTMAN J B.
XX (OKEE/) O'KEEFE T L.
XX (OZKA/) OZKAYNAK E.
XX (HEAL/) HEALEY J J.
XX
XX Rottman JB, O'Keefe TL, Ozkaynak E, Healey JJ;
XX
XX MPI: 2003-829643/77.
XX
XX N-PSDB; ADM64585.
DR

XX
PT New TANGO 197 and/or TANGO 216 fusion polypeptides for preventing or
PT ameliorating symptoms of anthrax comprises a von Willebrand factor A-like
PT domain (WVF) amino acid sequence and an amino acid sequence heterologous
PT to the WVF.
XX
XX
PS Claim 44; SEQ ID NO 10; 64pp; English.
XX
XX The invention describes a fusion polypeptide comprising a von Willebrand
XX factor A-like domain (WVF) amino acid sequence and an amino acid sequence
XX heterologous to the WVF. Also described are: a method of preventing or
XX ameliorating a symptom of anthrax in a subject thought to be at risk for
XX exposure to or suspected of having been exposed to Bacillus anthracis;
XX and a pharmaceutical composition comprising the novel fusion polypeptide.
XX The composition and method are useful in preventing or ameliorating
XX symptoms of cutaneous and/or inhalation anthrax. This is the amino acid
XX sequence of a fusion protein comprising human TANGO197 extracellular
XX region, minus the DG residues closest to the transmembrane region, and
XX immunoglobulin G (IgG) fragment of crystallisation (Fc) with mutations
XX L235A and G237A.
XX
SQ Sequence 384 AA;

Query Match 100.0%; Score 943; DB 7; Length 384;
Best Local Similarity 100.0%; Pred. No. 2.5e-96;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFDLYFLDKSGSVLHHMNEIYFVEQLAHKFI SPQLRMSFIVSTRGTLMLKLTEDREQ 60
DB 42 GFDLYFLDKSGSVLHHMNEIYFVEQLAHKFI SPQLRMSFIVSTRGTLMLKLTEDREQ 101
QY 61 IRGGLSELQKVLPGSDTYMHGFERASEQIYYENRQGYRTASVIALTDGELHEDLFFYS 120
DB 102 IRGGLSELQKVLPGSDTYMHGFERASEQIYYENRQGYRTASVIALTDGELHEDLFFYS 161
QY 121 EREANRSRDLGALVYCVGVDFNETQLARIADSKDHVPFVNDGFQALQGIHSLKKSCT 180
DB 162 EREANRSRDLGALVYCVGVDFNETQLARIADSKDHVPFVNDGFQALQGIHSLKKSCT 221
QY 181 E 181
DB 222 E 222

RESULT 14
AAE01439
ID AAE01439 standard; protein; 403 AA.
XX
AC AAE01439;
XX
XX
DT 17-JUL-2001 (first entry)
XX
XX Human gene 4 encoded secreted protein HMLPR02, SEQ ID NO:94.
XX
XX Human; secreted protein; proliferative disorder; cancer; tumour;
XX foetal abnormality; developmental abnormality; haematopoietic disorder;
XX immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
XX inflammation; allergy; neurological disorder; Alzheimer's disease;
XX Parkinson's disease; cognitive disorder; schizophrenia; asthma;
XX skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
XX cardiovascular disorder; angiogenic disorder; kidney disorder;
XX gastrointestinal disorder; pregnancy-related disorder;
XX endocrine disorder; infection; wound healing; vulnerability; cell culture;
XX chemotaxis; food additive; gene therapy; binding partner identification;
XX chromosome 19.
XX
XX Homo sapiens.
OS
FH Key
FH Peptide 1.27
FT /label=Signal_peptide
FT Protein 28.403
FT /note="Mature human secreted protein"

XX WO200134626-A1.
XX 17-MAY-2001.
XX 01-NOV-2000; 2000MO-US030045.
XX 05-NOV-1999; 99US-0163581P.
XX 30-JUN-2000; 2000US-0215133P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Ruben SM, Komatsu GA, Moore PA, Birze CE, Ni J;
XX WPI; 2001-308778/32.
XX N-PSDB; AAD05303.
XX New nucleic acid molecules encoding 28 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives.
XX Claim 11; Page 485-486; 562pp; English.
XX AAD05300-AAD05379 represent cDNAs corresponding to 28 human secreted
CC protein genes, and AAE01436-AAE01513 represent the proteins they encode.
CC AAE01514-AAE01544 represent human secreted protein fragments or variants.
CC The genes and their secreted proteins are useful for preventing, treating
CC or ameliorating medical conditions, e.g., by protein or gene therapy.
CC Pathological conditions can be diagnosed by determining the amount of the
CC new protein in a sample or by determining the presence of mutations in
CC the new genes. Specific uses are described for each of the 28 genes,
CC based on the tissues in which they are most highly expressed, and include
CC developing products for the diagnosis or treatment of proliferative
CC disorders, cancer, tumors, foetal and developmental abnormalities,
CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,
CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
CC angiogenic disorders, kidney disorders, gastrointestinal disorders,
CC pregnancy-related disorders, endocrine disorders, and infections. The
CC proteins can also be used to aid wound healing and epithelial cell
CC proliferation, to prevent skin aging due to sunburn, to maintain organs
CC before transplantation, for supporting cell culture of primary tissues,
CC to regenerate tissues, to identify their cognate ligands or binding
CC partners, and in chemotaxis, and can be used as a food additive or
CC preservative to modify storage properties. Antibodies specific for a
CC protein of the invention can be used in alleviating symptoms associated
CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,
CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The
CC present sequence represents a human secreted protein of the invention
XX
SQ Sequence 403 AA;
Query Match 100.0%; Score 943; DB 4; Length 403;
Best Local Similarity 100.0%; Pred. No. 2, 7e-96;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GFDLYFLIDKSGSVLHNMNEIYFVEQLAKHFI SPQLRMSFIYVSTGTTLMKLTEDREQ 60
DB 42 GFDLYFLIDKSGSVLHNMNEIYFVEQLAKHFI SPQLRMSFIYVSTGTTLMKLTEDREQ 101
QY 61 IROGLEELQKVLPGGDTYMHGEPFASQIYENRQGRYASVITALTDELHEDLFPYS 120
DB 102 IROGLEELQKVLPGGDTYMHGEPFASQIYENRQGRYASVITALTDELHEDLFPYS 161
QY 121 EREANRSDIGAIYVCVGDQFNETQLARLADSKDHVPVNDGFQALOGIHSILKXSCI 180
DB 162 EREANRSDIGAIYVCVGDQFNETQLARLADSKDHVPVNDGFQALOGIHSILKXSCI 221
QY 181 E 181
DB 222 E 222

RESULT 15
ABG63874
ID ABG63874 standard; protein; 403 AA.
XX AC ABG63874;
XX 27-AUG-2002 (first entry)
XX DE Human albumin fusion protein #549.
XX KW Albumin fusion protein; therapeutic protein X; human albumin; HA;
KW human serum albumin; HSA; cancer; reproductive disorder;
KW digestive disorder; immune disorder; endocrine disorder;
KW haematopoietic disorder; neural disorder; connective disorder;
KW cyrostatic; antiinfectility; antiinflammatory; anticancer;
KW immunomodulator; anti-HIV; antidiabetic; haemostatic; neurotropic;
KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
KW osteopathic; antiarthritic.
XX
XX Homo sapiens.
OS Synthetic.
OS
XX WO200177137-A1.
XX 18-OCT-2001.
XX 12-APR-2001; 2001MO-US011988.
XX 12-APR-2000; 2000US-0229358P.
XX 25-APR-2000; 2000US-0199384P.
XX 21-DEC-2000; 2000US-0256931P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Haseltine WA;
PI WPI; 2002-010886/01.
XX
XX New fusion protein for treating disease e.g. diabetes comprises an
PT albumin fused to a therapeutic protein.
XX
XX Claim 1; Page 874-875; 2102pp; English.
XX The present invention relates to albumin fusion proteins comprising a
CC therapeutic protein X and human albumin (HA, also known as human serum
CC albumin, HSA). The proteins are useful for treating a disease or disorder
CC that may be modulated by therapeutic protein X. The albumin extends the
CC shelf-life of protein X, and may increase its biological in vitro/in vivo
CC activity. The protein is useful for treating and diagnosing disorders
CC such as cancer, reproductive disorders, digestive disorders (e.g. Crohn's
CC disease, ulcerative colitis), immune disorders (e.g. acquired
CC immunodeficiency syndrome, AIDS), endocrine disorders (e.g. diabetes),
CC haematopoietic disorders, neural disorders (e.g. Alzheimer's,
CC Parkinson's, Creutzfeldt-Jacob disease, encephalomyelitis, meningitis,
CC schizophrenia), and connective disorders (e.g. osteoporosis, arthritis).
CC ABG63326-ABG65518 represent albumin fusion proteins of the invention
XX
SQ Sequence 403 AA;
Query Match 100.0%; Score 943; DB 5; Length 403;
Best Local Similarity 100.0%; Pred. No. 2, 7e-96;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GFDLYFLIDKSGSVLHNMNEIYFVEQLAKHFI SPQLRMSFIYVSTGTTLMKLTEDREQ 60
DB 42 GFDLYFLIDKSGSVLHNMNEIYFVEQLAKHFI SPQLRMSFIYVSTGTTLMKLTEDREQ 101
QY 61 IROGLEELQKVLPGGDTYMHGEPFASQIYENRQGRYASVITALTDELHEDLFPYS 120
DB 102 IROGLEELQKVLPGGDTYMHGEPFASQIYENRQGRYASVITALTDELHEDLFPYS 161

Qy	121	EREANRSRDIGAIYYCVGVDFNETOLARIADSKOHVFPVNDGFOALOGIIHSILKSCI	180
Db	162	EREANRSRDIGAIYYCVGVDFNETOLARIADSKOHVFPVNDGFOALOGIIHSILKSCI	221
Qy	181	E 181	
Db	222	E 222	

Search completed: June 13, 2005, 19:56:54
Job time : 68.0133 secs

OM protein - protein search, using sw model

Run on: June 13, 2005, 19:46:02 ; Search time 17,4748 Seconds
(without alignments)
773.200 Million cell updates/sec

Title:	US-09-970-076-2_COPY_42_222
Perfect score:	943
Sequence:	1 GFDLYFILDKSGSVLHMNE.....DFQALQGIHSILKSCIE 181

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

```

Post-processing: Minimum Match 0% Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents AA:*

```

1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfl1.pep:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB	ID	Description
1	141.5	15.0	1151	1	US-08-286-889-37
2	141.5	15.0	1151	1	US-08-485-618-37
3	141.5	15.0	1151	1	US-08-362-652-37
4	141.5	15.0	1151	1	US-08-605-672-37
5	141.5	15.0	1151	2	US-08-482-293A-37
6	141.5	15.0	1151	2	US-08-943-363-37
7	141.5	15.0	1151	3	US-09-193-043-37
8	141.5	15.0	1151	4	US-09-688-107A-37
9	141.5	15.0	1151	4	US-09-350-259-37
10	141.5	15.0	1161	1	US-08-485-618-55
11	141.5	15.0	1161	1	US-08-352-652-55
12	141.5	15.0	1161	2	US-08-605-672-55
13	141.5	15.0	1161	2	US-08-482-293A-55
14	141.5	15.0	1161	2	US-08-943-363-55
15	141.5	15.0	1161	3	US-09-193-043-55
16	141.5	15.0	1161	4	US-09-688-107A-55
17	141.5	15.0	1161	4	US-09-350-259-55
18	138.5	14.7	1155	1	US-08-485-618-46
19	138.5	14.7	1155	1	US-08-485-618-46
20	138.5	14.7	1155	1	US-08-362-652-46
21	138.5	14.7	1155	2	US-08-605-672-46
22	138.5	14.7	1155	2	US-08-482-293A-46
23	138.5	14.7	1155	3	US-08-943-363-46
24	138.5	14.7	1155	3	US-09-193-043-46
25	138.5	14.7	1155	4	US-09-688-107A-46
26	138.5	14.7	1155	4	US-09-350-259-46
27	138.5	14.7	1161	1	US-08-485-618-53

45	128.5	14.1	1153	2	US-08-605-672-3	Sequence 3, Appl1
44	132.5	14.1	1153	1	US-08-362-652-53	Sequence 53, Appl1
43	132.5	14.1	1153	1	US-08-605-672-53	Sequence 53, Appl1
42	132.5	14.1	1153	1	US-08-482-293A-53	Sequence 53, Appl1
41	132.5	14.1	1153	2	US-08-949-363-53	Sequence 53, Appl1
40	132.5	14.1	1152	6	US-09-193-043-53	Sequence 53, Appl1
39	132.5	14.1	1152	6	US-09-688-307A-53	Sequence 53, Appl1
38	132.5	14.1	1152	4	US-09-350-259-53	Sequence 53, Appl1
37	132.5	14.1	1152	2	US-09-795-872-5	Sequence 1, Appl1
36	132.5	14.1	1152	2	PCT-US95-0439-1	Sequence 1, Appl1
35	132.5	14.1	1152	2	US-08-476-062A-43	Sequence 43, Appl1
34	132.5	14.1	1152	5	PCT-US96-01314-43	Sequence 43, Appl1
33	132.5	14.1	1152	6	54243399-2	Patent No. 54243399-2
32	138.5	14.7	1151	3	US-08-173-497-3	Sequence 3, Appl1
31	138.5	14.7	1151	3	US-08-286-889-3	Sequence 3, Appl1
30	138.5	14.7	1151	3	US-08-485-618-3	Sequence 3, Appl1
29	138.5	14.7	1151	3	US-08-362-652-3	Sequence 3, Appl1
28	138.5	14.7	1151	1	US-08-605-672-3	Sequence 53, Appl1
27	138.5	14.7	1151	1	US-08-362-652-53	Sequence 53, Appl1
26	138.5	14.7	1151	2	US-08-605-672-53	Sequence 53, Appl1
25	138.5	14.7	1151	2	US-08-949-363-53	Sequence 53, Appl1
24	138.5	14.7	1151	3	US-09-193-043-53	Sequence 53, Appl1
23	138.5	14.7	1151	4	US-09-350-259-53	Sequence 53, Appl1
22	138.5	14.7	1151	4	US-09-795-872-5	Sequence 1, Appl1
21	138.5	14.7	1151	4	PCT-US95-0439-1	Sequence 1, Appl1
20	138.5	14.7	1151	4	US-08-476-062A-43	Sequence 43, Appl1
19	138.5	14.7	1151	5	PCT-US96-01314-43	Sequence 43, Appl1
18	138.5	14.7	1151	6	54243399-2	Patent No. 54243399-2
17	138.5	14.7	1151	6	US-08-173-497-3	Sequence 3, Appl1
16	138.5	14.7	1151	3	US-08-286-889-3	Sequence 3, Appl1
15	138.5	14.7	1151	3	US-08-485-618-3	Sequence 3, Appl1
14	138.5	14.7	1151	3	US-08-362-652-3	Sequence 3, Appl1
13	138.5	14.7	1151	1	US-08-605-672-3	Sequence 53, Appl1
12	138.5	14.7	1151	1	US-08-362-652-53	Sequence 53, Appl1
11	138.5	14.7	1151	2	US-08-605-672-53	Sequence 53, Appl1
10	138.5	14.7	1151	2	US-08-949-363-53	Sequence 53, Appl1
9	138.5	14.7	1151	3	US-09-193-043-53	Sequence 53, Appl1
8	138.5	14.7	1151	4	US-09-350-259-53	Sequence 53, Appl1
7	138.5	14.7	1151	4	US-09-795-872-5	Sequence 1, Appl1
6	138.5	14.7	1151	4	PCT-US95-0439-1	Sequence 1, Appl1
5	138.5	14.7	1151	4	US-08-476-062A-43	Sequence 43, Appl1
4	138.5	14.7	1151	5	PCT-US96-01314-43	Sequence 43, Appl1
3	138.5	14.7	1151	6	54243399-2	Patent No. 54243399-2
2	138.5	14.7	1151	6	US-08-173-497-3	Sequence 3, Appl1
1	138.5	14.7	1151	3	US-08-286-889-3	Sequence 3, Appl1

ALIGNMENTS

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Db 142 DIAFLIDSGSSINQRFQKMDPVKALMGEPASTSTLSLMQYNSILKHTFTFEFNIL 201
Qy 57 DREIIRQGLELEOKVLPFGDITYMEGFERASEQIYYENRGYRTA-SVIALTJGELHED 115
Db 202 DPOSIVDPVLOQ-----GLTYTATGIRTYMEELFHSKNGSRKSAKILLVITDGOXYRD 256
Qy 116 LFFYSE--REANRSRDGAIVYCVGYD-FNE-TQLARI-----ADSKDHVPYNDGFOA 166
Db 257 PLEYSVDVTPAADKA--GIRYALGVDAFOEPALKEIWTIGSAPPQDHVFKVGN-FAA 312
Qy 167 LOGIHSILKK 177
Db 313 LRSIORQLOEK 323

RESULT 2
US-08-485-618-37
; Sequence 37, Application US/08485618
; Patent No. 5728533
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vlieten, Monica
; TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,618
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32797
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1151 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-485-618-37

Query Match 15.0%; Score 141.5; DB 1; Length 1151;
Best Local Similarity 27.7%; Pred. No. 4.2e-07;
Matches 53; Conservative 38; Mismatches 75; Indels 25; Gaps 10;
Qy 3 DLYFLIDKSGSV-LHNNNEIYVFEOLAHKFIQPLMSFIVSTGTTLMKTE----- 56
Db 142 DIAFLIDSGSSINQRFQKMDPVKALMGEPASTSTLSLMQYNSILKHTFTFEFNIL 201

Qy 57 DREIIRQGLELEOKVLPFGDITYMEGFERASEQIYYENRGYRTA-SVIALTJGELHED 115
Db 202 DPOSIVDPVLOQ-----GLTYTATGIRTYMEELFHSKNGSRKSAKILLVITDGOXYRD 256
Qy 116 LFFYSE--REANRSRDGAIVYCVGYD-FNE-TQLARI-----ADSKDHVPYNDGFOA 166
Db 257 PLEYSVDVTPAADKA--GIRYALGVDAFOEPALKEIWTIGSAPPQDHVFKVGN-FAA 312
Qy 167 LOGIHSILKK 177
Db 313 LRSIORQLOEK 323

RESULT 3
US-08-362-652-37
; Sequence 37, Application US/08362652
; Patent No. 5766850
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vlieten, Monica
; TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,652
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32391
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1151 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-362-652-37

Query Match 15.0%; Score 141.5; DB 1; Length 1151;
Best Local Similarity 27.7%; Pred. No. 4.2e-07;
Matches 53; Conservative 38; Mismatches 75; Indels 25; Gaps 10;
Qy 3 DLYFLIDKSGSV-LHNNNEIYVFEOLAHKFIQPLMSFIVSTGTTLMKTE----- 56
Db 142 DIAFLIDSGSSINQRFQKMDPVKALMGEPASTSTLSLMQYNSILKHTFTFEFNIL 201
Qy 57 DREIIRQGLELEOKVLPFGDITYMEGFERASEQIYYENRGYRTA-SVIALTJGELHED 115
Db 202 DPOSIVDPVLOQ-----GLTYTATGIRTYMEELFHSKNGSRKSAKILLVITDGOXYRD 256
Qy 116 LFFYSE--REANRSRDGAIVYCVGYD-FNE-TQLARI-----ADSKDHVPYNDGFOA 166

Db 257 PLEYSVDIIPADKA---GIIRYAI GVD AFOEPALKE LNTIGSAPQDHVFKGN-FAA 312
QY 167 LOGIHSILKK 177
Db 313 LRSIORQLQEK 323

RESULT 4

US-08-605-672-37
Sequence 37, Application US/08605672
Patent No. 5817515
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vlieten, Monica
TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/605,672
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 1151 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-605-672-37

Query Match 15.0%; Score 141.5; DB 2; Length 1151;
Best Local Similarity 27.7%; Pred. No. 4.2e-07;
Matches 53; Conservative 38; Mismatches 75; Indels 25; Gaps 10;

QY 3 DLYEILDKSGSV-LHMHNEIYYFVEOLAHKRISPOLMSFIVFSTRTGTLMLKTE----- 56
Db 142 DIAFLIGSSGSSINORDPAQKMDFKALMGEPASTLFSLMQYSNLIKTHFTFEFKNIL 201
QY 57 DREQIRGSELQKLVLPFGSDTYHMEGPERASEQIYYENRQGYRTA-SVIALTLTGELHED 115
Db 202 DPGSLVDPIVQLQ-----GLYTTATGIRTWBELFHSKNGSRKSAKKILLVITDGQKRD 256
QY 116 LFFYSE--REANRGRDGAIVYCVGVD-FNE-TQLARI-----ADSKDHVPVNDGFOA 166
Db 257 PLEYSVDIIPADKA---GIIRYAI GVD AFOEPALKE LNTIGSAPQDHVFKGN-FAA 312

QY 167 LOGIHSILKK 177
Db 313 LRSIORQLQEK 323

RESULT 5

US-08-482-293A-37
Sequence 37, Application US/08482293A
Patent No. 5831029
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vlieten, Monica
TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,293A
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 1151 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-482-293A-37

Query Match 15.0%; Score 141.5; DB 2; Length 1151;
Best Local Similarity 27.7%; Pred. No. 4.2e-07;
Matches 53; Conservative 38; Mismatches 75; Indels 25; Gaps 10;

QY 3 DLYEILDKSGSV-LHMHNEIYYFVEOLAHKRISPOLMSFIVFSTRTGTLMLKTE----- 56
Db 142 DIAFLIGSSGSSINORDPAQKMDFKALMGEPASTLFSLMQYSNLIKTHFTFEFKNIL 201
QY 57 DREQIRGSELQKLVLPFGSDTYHMEGPERASEQIYYENRQGYRTA-SVIALTLTGELHED 115
Db 202 DPGSLVDPIVQLQ-----GLYTTATGIRTWBELFHSKNGSRKSAKKILLVITDGQKRD 256
QY 116 LFFYSE--REANRGRDGAIVYCVGVD-FNE-TQLARI-----ADSKDHVPVNDGFOA 166
Db 257 PLEYSVDIIPADKA---GIIRYAI GVD AFOEPALKE LNTIGSAPQDHVFKGN-FAA 312
QY 167 LOGIHSILKK 177

Db 313 LRSIQQLQEK 323

```

1      RESULT 6
2      US-08-943-363-37
3      Sequence 37, Application US/08943363
4      Patent No. 5837478
5      GENERAL INFORMATION:
6      APPLICANT: Gallatin, W. Michael
7      APPLICANT: Van der Vieren, Monica
8      TITLE OF INVENTION: No. 5837478e1 Human 2 Integrin Alpha Subunit
9      NUMBER OF SEQUENCES: 114
10     CORRESPONDENCE ADDRESS:
11     ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borum
12     STREET: 233 South Wacker Drive, 6300 Sear Tower
13     CITY: Chicago
14     STATE: Illinois
15     COUNTRY: United States
16     ZIP: 60606-6402
17     COMPUTER READABLE FORM:
18     MEDIUM TYPE: Floppy disk
19     COMPUTER: IBM PC compatible
20     OPERATING SYSTEM: PC-DOS/MS-DOS
21     SOFTWARE: Patent In Release #1.0, Version #1.25
22     CURRENT APPLICATION DATA:
23     APPLICATION NUMBER: US/08/943.363
24     FILING DATE:
25     CLASSIFICATION: 530
26     PRIOR APPLICATION DATA:
27     APPLICATION NUMBER: US 08/173.497
28     FILING DATE: 23-DEC-1993
29     PRIOR APPLICATION DATA:
30     APPLICATION NUMBER: US 08/286,889
31     FILING DATE: 5-AUG-1994
32     PRIOR APPLICATION DATA:
33     APPLICATION NUMBER: US 08/362,652
34     FILING DATE: 21-DEC-1994
35     ATTORNEY/AGENT INFORMATION:
36     NAME: Williams Jr., Joseph A.
37     REGISTRATION NUMBER: 38,659
38     REFERENCE/DOCKET NUMBER: 27866/32664
39     TELECOMMUNICATION INFORMATION:
40     TELEPHONE: 312-474-6300
41     TELEFAX: 312-474-0448
42     TELEX: 25-3856
43     INFORMATION FOR SEQ ID NO: 37:
44     SEQUENCE CHARACTERISTICS:
45     LENGTH: 1151 amino acids
46     TYPE: amino acid
47     TOPOLOGY: linear
48     MOLECULE TYPE: protein
49     IS-08-943-363-37

```

```

RESULT 7
US-09-193-043-37
; Sequence 37, Application US/09193043
; Patent No. 6251395
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vlieten, Monica
; TITLE OF INVENTION: No. 6251395el Human 2
; FILE REFERENCE: 27666/35004
; CURRENT APPLICATION NUMBER: US/09/193, 043
; CURRENT FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173, 497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286, 889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362, 652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943, 363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 37
; LENGTH: 1151
; TYPE: prt
; ORGANISM: Rattus rattus
US-09-193-043-37

```

RESULT 8
US-09-688-307A-37
Sequence 37, Application US/09688307A
Patent No. 6432404
GENERAL INFORMATION:
APPLICANT: Gallatin, Michael W.
APPLICANT: Van der Vlieten, Monica
TITLE OF INVENTION: No. 6432404e1 Human Beta-2-
FILE REFERENCE: 27866/36646
CURRENT APPLICATION NUMBER: US/09/688,307A
CURRENT FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 09/123,043
PRIOR FILING DATE: 1998-11-16
PRIOR APPLICATION NUMBER: 08/605,672
PRIOR FILING DATE: 1996-02-22
PRIOR APPLICATION NUMBER: 08/173,497
PRIOR FILING DATE: 1993-12-23
PRIOR APPLICATION NUMBER: 08/286,889
PRIOR FILING DATE: 1994-08-05
PRIOR APPLICATION NUMBER: 08/352,652
PRIOR FILING DATE: 1994-12-21
PRIOR APPLICATION NUMBER: 08/943,363
PRIOR FILING DATE: 1997-10-03

```

/ NUMBER OF SEQ ID NOS: 114
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 37
/ LENGTH: 1151
/ TYPE: PRT
/ ORGANISM: Rattus rattus
/ FEATURES:
/   NAME/KEY: misc_feature
/   LOCATION: 361
/   OTHER INFORMATION: Xaa = any or unknown amino acid
/   NAME/KEY: misc_feature
/   LOCATION: 464
/   OTHER INFORMATION: Xaa = any or unknown amino acid
/   NAME/KEY: misc_feature
/   LOCATION: 486
/   OTHER INFORMATION: Xaa = any or unknown amino acid
/   NAME/KEY: misc_feature
/   LOCATION: 506
/   OTHER INFORMATION: Xaa = any or unknown amino acid
/   NAME/KEY: misc_feature
/   LOCATION: 1117
/   OTHER INFORMATION: Xaa = any or unknown amino acid
/   NAME/KEY: misc_feature
/   LOCATION: 1118
/   OTHER INFORMATION: Xaa = any or unknown amino acid
/ US-09-688-307A-37

```

```

Query Match          15.0%; Score 141.5; DB 4; Length 1151;
Best Local Similarity 27.7%; Pred. No. 4.2e-07;
Matches 53; Conservative 38; Mismatches 75; Indels 25; Gaps 10;

```

```

Qy 3 DLYEILDKSGSV-LHNNHEIYFVEQLAHKFIQPMSPFVSTRTGLMKLTE----- 56
Db 142 DIAFLDSSGSSINGRDPAQKMDFKALMGEPASTLPSLMQYSNLKHTFTPEFNKL 201
Qy 57 DREQIRQGLBELQKVLPGDPTMHGEPFASSEQIYYENRQGYRTA-SVIALTDGELHED 115
Db 202 DPOSIVPIVOLQ-----GLTYATGIRTYWELFHSKNGSRKSAKILLVITDGOKYRD 256
Qy 116 LFPFSE--REANSRDIAIVYCVGYVD-FNE-TQLARI-----ADSKDHFPVNDGFOA 166
Db 257 PLEISDVIAPADKA--GIIRYALGVGDAPQEPALKEIWTIGSAPPQDHVKYGN-FAA 312
Qy 167 LQGIHSILKK 177
Db 313 LRSTGRQLQEK 323

```

```

RESULT 9
US-09-350-259-37
/ Sequence 37, Application US/09350259
/ Patent No. 6620915
/ GENERAL INFORMATION:
/   APPLICANT: Gallatin, Michael W.
/   TITLE OF INVENTION: No. 6620915el Human 2
/   FILE REFERENCE: 27866/35004
/   CURRENT APPLICATION NUMBER: US/09/350,259
/   FILING DATE: 1999-07-08
/   EARLIER APPLICATION NUMBER: 09/193,043
/   EARLIER FILING DATE: 1998-11-16
/   EARLIER APPLICATION NUMBER: 08/173,497
/   EARLIER FILING DATE: 1993-12-23
/   EARLIER APPLICATION NUMBER: 08/286,889
/   EARLIER FILING DATE: 1994-08-05
/   EARLIER APPLICATION NUMBER: 08/362,652
/   EARLIER FILING DATE: 1994-12-21
/   EARLIER APPLICATION NUMBER: 08/943,363
/   EARLIER FILING DATE: 1997-10-03
/   NUMBER OF SEQ ID NOS: 114
/   SOFTWARE: PatentIn Ver. 2.0
/   SEQ ID NO 37
/   LENGTH: 1151

```

```

/ TYPE: PRT
/ ORGANISM: Rattus rattus
/ US-09-350-259-37

```

```

Query Match          15.0%; Score 141.5; DB 4; Length 1151;
Best Local Similarity 27.7%; Pred. No. 4.2e-07;
Matches 53; Conservative 38; Mismatches 75; Indels 25; Gaps 10;

```

```

Qy 3 DLYEILDKSGSV-LHNNHEIYFVEQLAHKFIQPMSPFVSTRTGLMKLTE----- 56
Db 142 DIAFLDSSGSSINGRDPAQKMDFKALMGEPASTLPSLMQYSNLKHTFTPEFNKL 201
Qy 57 DREQIRQGLBELQKVLPGDPTMHGEPFASSEQIYYENRQGYRTA-SVIALTDGELHED 115
Db 202 DPOSIVPIVOLQ-----GLTYATGIRTYWELFHSKNGSRKSAKILLVITDGOKYRD 256
Qy 116 LFPFSE--REANSRDIAIVYCVGYVD-FNE-TQLARI-----ADSKDHFPVNDGFOA 166
Db 257 PLEISDVIAPADKA--GIIRYALGVGDAPQEPALKEIWTIGSAPPQDHVKYGN-FAA 312
Qy 167 LQGIHSILKK 177
Db 313 LRSTGRQLQEK 323

```

```

RESULT 10
US-08-485-618-55
/ Sequence 55, Application US/08485618
/ Patent No. 5728533
/ GENERAL INFORMATION:
/   APPLICANT: Gallatin, W. Michael
/   TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit
/   NUMBER OF SEQUENCES: 103
/   CORRESPONDENCE ADDRESS:
/     ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
/     STREET: 233 South Wacker Drive, 6300 Sear Tower
/     CITY: Chicago
/     STATE: Illinois
/     COUNTRY: United States
/     ZIP: 60606-6402
/   COMPUTER READABLE FORM:
/     MEDIUM TYPE: Floppy disk
/     OPERATING SYSTEM: PC-DOS/MS-DOS
/     SOFTWARE: PatentIn Release #1.0, Version #1.25
/   CURRENT APPLICATION DATA:
/     APPLICATION NUMBER: US/08/485,618
/     FILING DATE:
/     CLASSIFICATION: 435
/     PRIOR APPLICATION DATA:
/       APPLICATION NUMBER: US 08/173,497
/       FILING DATE: 23-DEC-1993
/       PRIOR APPLICATION DATA:
/         APPLICATION NUMBER: US 08/286,889
/         FILING DATE: 5-AUG-1994
/         PRIOR APPLICATION DATA:
/           APPLICATION NUMBER: US 08/362,652
/           FILING DATE: 21-DEC-1994
/           ATTORNEY/AGENT INFORMATION:
/             NAME: Williams Jr., Joseph A.
/             REGISTRATION NUMBER: 38,659
/             REFERENCE/DOCKET NUMBER: 27866/32797
/             TELECOMMUNICATION INFORMATION:
/               TELEPHONE: 312-474-6300
/               TELEFAX: 312-474-0448
/               TELEX: 25-3856
/   INFORMATION FOR SEQ ID NO: 55:
/     SEQUENCE CHARACTERISTICS:
/       LENGTH: 1161 amino acids
/       TYPE: amino acid
/       TOPOLOGY: linear
/     MOLECULE TYPE: protein

```


Db 152 DIAPLDGSSGINSQNDPFAQMKDPVAKALMBEPASTSTLPSLMQYSNLKHFHTTERKNLL 211

Qy 57 DREGIROGLEBIQKVLPGGDTYMHGFEFPASEQIYENKQGYRTA-SVIALTDEGLHED 115

Db 212 DPQSLVDPYIVQLQ-----GLTTATGIRKVMBELPFSKNGSRKSAGKILLVITDQOKYRD 266

Qy 116 LFPYSE--REANRSHDLGAIIVYCVGKVD-FNE-TQLARI-----ADSKDHVPFVNDGFOA 166

Db 267 PLEYSDDVLPADKA---GIRRAIGVGDAFQEPALTAKELNTIGSAPPDHVFKNVGN-FAA 322

Qy 167 LQGIHHSILKK 177

Db 323 LRSTORQLQEK 333

```

1      RESULT 13
2      US-08-482-293A-55
3      Sequence 55, Application US/08482293A
4      Patent No. 5831029
5      GENERAL INFORMATION:
6      APPLICANT: Gallatin, W. Michael
7      APPLICANT: Van der Vieren, Monica
8      TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit
9      NUMBER OF SEQUENCES: 103
10     CORRESPONDENCE ADDRESS:
11     ADDRESSEE: Marshhall, O'Toole, Gerstein, Murray & Borun
12     STREET: 233 South Wacker Drive, 6300 Sear Tower
13     CITY: Chicago
14     STATE: Illinois
15     COUNTRY: United States
16     ZIP: 60606-6402
17     COMPUTER READABLE FORM:
18     MEDIUM TYPE: Floppy disk
19     COMPUTER: IBM PC compatible
20     OPERATING SYSTEM: PC-DOS/MS-DOS
21     SOFTWARE: Patent Release #1.0, Version #1.25
22     CURRENT APPLICATION DATA:
23     APPLICATION NUMBER: US/08/482,293A
24     FILING DATE:
25     CLASSIFICATION: 530
26     PRIOR APPLICATION DATA:
27     APPLICATION NUMBER: US 08/173,497
28     FILING DATE: 23-DEC-1993
29     PRIOR APPLICATION DATA:
30     APPLICATION NUMBER: US 08/286,889
31     FILING DATE: 5-AUG-1994
32     PRIOR APPLICATION DATA:
33     APPLICATION NUMBER: US 08/362,652
34     FILING DATE: 21-DEC-1994
35     ATTORNEY/AGENT INFORMATION:
36     NAME: Williams Jr., Joseph A.
37     REGISTRATION NUMBER: 38,659
38     REFERENCE/DOCKET NUMBER: 27866/32684
39     TELECOMMUNICATION INFORMATION:
40     TELEPHONE: 312-474-6300
41     TELEFAX: 312-474-0448
42     TELEX: 25-3856
43     INFORMATION FOR SEQ ID NO: 55:
44     SEQUENCE CHARACTERISTICS:
45     LENGTH: 1161 amino acids
46     TYPE: amino acid
47     TOPOLOGY: linear
48     MOLECULE TYPE: protein
49     US-08-482-293A-55
50
51     Query Match      15.0%; Score 141.5; DB 2; Length 1161;
52     Best Local Similarity 27.7%; Pred. No. 4.2e-07;
53     Matches 53; Conservative 38; Mismatches 75; Indels 25; Gaps 10;
54
55     3 DLYVILDKSGGV-LHHMNEIYYFVQLAHKFIISPOLAKSFIVFSRGTTLNKLTE----- 56
56     | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
57     152 DIAPLIDSSGGINORDPQMDPVVALMGEPASTSTLPSLMQVSYILKTHFTTFEKNIL 211

```

```

QY 57 REQDPTMGLEBELOKLPGDPTVMHGFELASQIYYENNGVRLA-SVIALLDGELHED 115
Db 212 DPOSIVLDEIVOLQ-----GLTYTANGIRTVMELFESKNGSRSAKKILLVITDGOXYRD 266
QY 116 LFPYSE--REANRSADLGAIYVCVGKD-FNE-TOLARI-----ADSKDHVFPVNDGFOA 166
Db 267 PLERSDVIYPADKA---GIIRYATGVGAFOEPTALKEINTIGSAPPDHFVKVGN-FAA 322
QY 167 LOGIITHSIUKK 177
Db 323 LRSIDRQLQLEK 333

```

```

01 RESULT 14
02 US-08-943-363-55
03 ; Sequence 55, Application US/08943363
04 ; Patent No 5837478
05 ; GENERAL INFORMATION:
06 APPLICANT: Gallatin, W. Michael
07 APPLICANT: Van der Vieren, Monica
08 TITLE OF INVENTION: No. 5837478el Human 2 Integrin Alpha Subunit
09 NUMBER OF SEQUENCES: 114
10 CORRESPONDENCE ADDRESS:
11 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
12 STREET: 233 South Wacker Drive, 6300 Sear Tower
13 CITY: Chicago
14 STATE: Illinois
15 COUNTRY: United States
16 ZIP: 60606-6402
17 COMPUTER READABLE FORM:
18 MEDIUM TYPE: Floppy disk
19 COMPUTER: IBM PC compatible
20 OPERATING SYSTEM: PC-DOS/MS-DOS
21 SOFTWARE: Patentin Release #1.0, Version #1.25
22 CURRENT APPLICATION DATA:
23 APPLICATION NUMBER: US/08/943,363
24 FILING DATE:
25 CLASSIFICATION: 530
26 PRIOR APPLICATION DATA:
27 APPLICATION NUMBER: US 08/173,497
28 FILING DATE: 23-DEC-1993
29 PRIOR APPLICATION DATA:
30 APPLICATION NUMBER: US 08/286,889
31 FILING DATE: 5-AUG-1994
32 PRIOR APPLICATION DATA:
33 APPLICATION NUMBER: US 08/362,652
34 FILING DATE: 21-DEC-1994
35 ATTORNEY/AGENT INFORMATION:
36 NAME: Williams Jr., Joseph A.
37 REGISTRATION NUMBER: 38,659
38 REFERENCE/DOCKET NUMBER: 27866/32684
39 TELECOMMUNICATION INFORMATION:
40 TELEPHONE: 312-474-6300
41 TELEFAX: 312-474-0448
42 TELEX: 25-3856
43 INFORMATION FOR SEQ ID NO: 55:
44 SEQUENCE CHARACTERISTICS:
45 LENGTH: 1161 amino acids
46 TYPE: amino acid
47 TOPOLOGY: linear
48 MOLECULE TYPE: protein
49 ;
50 US-08-943-363-55
51
52 Query Match 15.0%; Score 141.5; DB 2; Length 1161;
53 Best Local Similarity 27.7%; Pred. No. 4.2e-07;
54 Matches 53; Conservative 38; Mismatches 75; Indels 25; Gaps 10
55
56 3 DLYFLIDSGSV-LHHNMRIYFYEQALHKRISPOLRRSFIIVSTRGTLTKLT-----56
57 |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
58 152 DIATLIDSSGINQDRPAQMDFVKALMGFEASITSLFSLMQSYILKTHPTFTFKNLL 211
59
60 57 DRDQIRGLEELQKVLPGGDITYMEGFERRASDQIYYENRGRTA-SVIALTDGLHED 115

```

```
Db      212 DPGSLVDPPIVQLQ-----GLTYTATGIRTWMEELFHSKNGSRKSAKKILIVITTDGQKYRD 266
QY      116 LFFYSE--REANRSRDIGAIYVCVVKD-FNE-TQLARI-----ADSKDHVPVNDGFOA 166
Db      267 PLEYSDEVIPADKA---GIIRYAIGVGDAFOEPTALKEIANTIGSAPPQDHVFKVGN-FAA 322
QY      167 LOGIHSILKK 177
Db      323 LRSIQRLQEK 333
```

RESULT 15

```
US-09-193-043-55
; Sequence 55, Application US/09193043
; Patent No. 6251395
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6251395e1 Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/193, 043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173, 497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286, 889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362, 652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943, 363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patentn Ver. 2.0
; SEQ ID NO 55
; LENGTH: 1161
; TYPE: PRT
; ORGANISM: Rattus rattus
US-09-193-043-55
```

```
Query Match      15.0%; Score 141.5; DB 3; Length 1161;
Best Local Similarity 27.7%; Pred. No. 4.2e-07;
Matches 53; Conservative 38; Mismatches 75; Indels 25; Gaps 10;
```

```
QY      3 DLYFIIDKSGSV-LHMHNEIYFVEQIAHKEFISPOLMSFIVSTRGTTMLKITE----- 56
Db      152 DIAFLIDSGSINQRDPQMKDFVKALMGEPASTLFLSMQYSNLIKTHFTTFEKNIL 211
QY      57 DREQIRQGLEELQKVLPGDPTVMHEGPERASEQIYENRQGYRTA-SVITALTGDELHED 115
Db      212 DPGSLVDPPIVQLQ-----GLTYTATGIRTWMEELFHSKNGSRKSAKKILIVITTDGQKYRD 266
QY      116 LFFYSE--REANRSRDIGAIYVCVVKD-FNE-TQLARI-----ADSKDHVPVNDGFOA 166
Db      267 PLEYSDEVIPADKA---GIIRYAIGVGDAFOEPTALKEIANTIGSAPPQDHVFKVGN-FAA 322
QY      167 LOGIHSILKK 177
Db      323 LRSIQRLQEK 333
```

Search completed: June 13, 2005, 20:05:10
Job time : 18.6748 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 20:03:29 ; Search time 58.0354 Seconds
(without alignments)
1195.537 Million cell updates/sec

Title: US-09-970-076-2_COPY_42_222

Perfect score: 943
Sequence: 1 GFDLYFLIDKSGSVLHMHNE.....DGFQALGGIHSILKSCIE 181

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1710399 seqs, 38333425 residues

Total number of hits satisfying chosen parameters: 1710399

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubppaa/PCTU_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubppaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubppaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubppaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubppaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubppaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubppaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/1/pubppaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	943	100.0	328	14	US-10-038-307-26 Sequence 26, Appl
2	943	100.0	328	14	US-10-201-292-26 Sequence 26, Appl
3	943	100.0	333	10	US-09-796-753-12 Sequence 12, Appl
4	943	100.0	333	14	US-10-038-307-2 Sequence 2, Appl1
5	943	100.0	333	14	US-10-201-292-2 Sequence 2, Appl1
6	943	100.0	342	14	US-10-038-307-22 Sequence 22, Appl
7	943	100.0	342	14	US-10-201-292-22 Sequence 22, Appl
8	943	100.0	345	14	US-10-038-307-24 Sequence 24, Appl
9	943	100.0	345	14	US-10-201-292-24 Sequence 24, Appl
10	943	100.0	403	11	US-09-833-245-621 Sequence 621, Appl
11	943	100.0	460	14	US-10-201-292-28 Sequence 28, Appl

12	943	100.0	479	14	US-10-201-292-32 Sequence 32, Appl
13	943	100.0	504	14	US-10-201-292-34 Sequence 34, Appl
14	943	100.0	529	14	US-10-201-292-36 Sequence 36, Appl
15	943	100.0	551	14	US-10-038-307-18 Sequence 18, Appl
16	943	100.0	551	14	US-10-201-292-18 Sequence 18, Appl
17	943	100.0	564	10	US-09-818-715-187 Sequence 187, Appl
18	943	100.0	564	10	US-09-818-715-232 Sequence 232, Appl
19	943	100.0	564	14	US-10-038-307-20 Sequence 20, Appl
20	943	100.0	564	14	US-10-201-292-20 Sequence 20, Appl
21	943	100.0	564	14	US-10-301-822-199 Sequence 199, Appl
22	943	100.0	564	16	US-10-408-765A-1823 Sequence 1823, Appl
23	943	100.0	564	16	US-10-474-794-187 Sequence 187, Appl
24	943	100.0	564	16	US-10-474-794-232 Sequence 232, Appl
25	938	99.5	403	11	US-09-833-245-620 Sequence 620, Appl
26	934	99.0	562	10	US-09-818-715-194 Sequence 194, Appl
27	934	99.0	562	10	US-09-818-715-301 Sequence 301, Appl
28	934	99.0	562	16	US-10-474-794-194 Sequence 194, Appl
29	934	99.0	562	16	US-10-474-794-301 Sequence 301, Appl
30	933	98.9	460	14	US-10-201-292-30 Sequence 30, Appl
31	922	97.8	534	14	US-10-038-307-12 Sequence 12, Appl
32	922	97.8	534	14	US-10-201-292-12 Sequence 12, Appl
33	922	97.8	543	14	US-10-038-307-10 Sequence 10, Appl
34	922	97.8	543	14	US-10-038-307-14 Sequence 14, Appl
35	922	97.8	543	14	US-10-038-307-16 Sequence 16, Appl
36	922	97.8	543	14	US-10-201-292-10 Sequence 10, Appl
37	922	97.8	543	14	US-10-201-292-14 Sequence 14, Appl
38	922	97.8	543	14	US-10-201-292-16 Sequence 16, Appl
39	900	95.4	538	13	US-10-047-542-99 Sequence 99, Appl
40	542	57.5	488	10	US-10-038-307-6 Sequence 6, Appl1
41	542	57.5	488	14	US-10-201-292-6 Sequence 6, Appl1
42	542	57.5	488	14	US-10-368-087-16 Sequence 16, Appl
43	542	57.5	488	15	US-10-104-047-2639 Sequence 2639, Appl
44	542	57.5	488	15	US-10-104-047-2639 Sequence 2639, Appl
45	540	57.3	487	10	US-09-796-753-54 Sequence 54, Appl

ALIGNMENTS

RESULT 1				
US-10-038-307-26				
Sequence 26, Application US/10038307				
Publication No. US20030134786A1				
GENERAL INFORMATION:				
APPLICANT: James B. ROTTMAN				
APPLICANT: Theresa L. O'KEEFE				
APPLICANT: Engin OZKANAK				
TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods				
FILE REFERENCE: 7853-253-999				
CURRENT APPLICATION NUMBER: US/10/038.307				
CURRENT FILING DATE: 2002-06-28				
NUMBER OF SEQ ID NOS: 26				
SOFTWARE: FastSeq for Windows Version 4.0				
SEQ ID NO 26				
LENGTH: 328				
TYPE: PRT				
ORGANISM: Homo sapiens				
US-10-038-307-26				
Query Match	100.0%	Score 943;	DB 14;	Length 328;
Best Local Similarity	100.0%;	Pred. No. 1.6e-90;		
Matches 181;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 1 GFDLYFLIDKSGSVLHMHNEIYVEQLAKHFIISPOLRMSFTVSTGTITMKLTEDREQ 60				
Db 42 GFDLYFLIDKSGSVLHMHNEIYVEQLAKHFIISPOLRMSFTVSTGTITMKLTEDREQ 101				
QY 61 IROGLEELQKVLPGSGDYYMHGFERASEQIYENRQGYRTASVITATDGLHDDLPFFYS 120				
Db 102 IROGLEELQKVLPGSGDYYMHGFERASEQIYENRQGYRTASVITATDGLHDDLPFFYS 161				
QY 121 EREANRSRDIGAIYVCVVDNFNETQARIADSKDHVFPVNDGFQALGGIHSILKSCIE 180				

```
Db      162 EREANRSDIGAIYCVGVDPFNETOLARIADSKDHVPVNDGFQALOGIIHSLKKSCT 221
      181 E 181
      222 E 222

RESULT 2
US-10-201-292-26
; Sequence 26, Application US/10201292
; Publication No. US20030144193A1
; GENERAL INFORMATION:
; APPLICANT: James B. ROTTMAN
; APPLICANT: Theresa L. O'KEEFE
; APPLICANT: Englin OZKAYNAK
; APPLICANT: Judith J. HEALEY
; TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
; FILE REFERENCE: 7853-253-999
; CURRENT APPLICATION NUMBER: US/10/201,292
; CURRENT FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-201-292-26

Query Match      100.0%; Score 943; DB 14; Length 328;
Best Local Similarity 100.0%; Pred. No. 1,6e-90;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GFDLYFLDKSGSVLHMHNEIYFVEQLAHKFTSPQLRMSFIVFSTRTTLMKLTEDREQ 60
      42 GFDLYFLDKSGSVLHMHNEIYFVEQLAHKFTSPQLRMSFIVFSTRTTLMKLTEDREQ 101
Db      61 IRGLEELQVLPFGDTYMHGFEFRASEQIYYENRQGYRRASVTIALTDGLHEDLFFYS 120
      102 IRGLEELQVLPFGDTYMHGFEFRASEQIYYENRQGYRRASVTIALTDGLHEDLFFYS 161
Qy      121 EREANRSDIGAIYCVGVDPFNETOLARIADSKDHVPVNDGFQALOGIIHSLKKSCT 180
      162 EREANRSDIGAIYCVGVDPFNETOLARIADSKDHVPVNDGFQALOGIIHSLKKSCT 221
Db      181 E 181
      222 E 222

RESULT 3
US-09-796-753-12
; Sequence 12, Application US/09796753
; Publication No. US20030027998A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-227-999
; CURRENT APPLICATION NUMBER: US/09/796,753
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 09/183,175
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 09/223,094
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/223,546
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/224,246
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/259,388
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 60/122,458
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 09/312,359
```

```
      PRIOR FILING DATE: 1999-05-14
      PRIOR APPLICATION NUMBER: 09/336,536
      PRIOR FILING DATE: 1999-06-18
      PRIOR APPLICATION NUMBER: 09/342,687
      PRIOR FILING DATE: 1999-06-29
      PRIOR APPLICATION NUMBER: 09/345,464
      PRIOR FILING DATE: 1999-06-30
      PRIOR APPLICATION NUMBER: 09/365,164
      PRIOR FILING DATE: 1999-07-30
      PRIOR APPLICATION NUMBER: 09/399,723
      PRIOR FILING DATE: 1999-09-20
      PRIOR APPLICATION NUMBER: 09/409,634
      PRIOR FILING DATE: 1999-09-30
      PRIOR APPLICATION NUMBER: 09/471,179
      PRIOR FILING DATE: 1999-12-23
      PRIOR APPLICATION NUMBER: 09/474,071
      PRIOR FILING DATE: 1999-12-29
      PRIOR APPLICATION NUMBER: 09/474,072
      PRIOR FILING DATE: 1999-12-29
      PRIOR APPLICATION NUMBER: 09/514,010
      PRIOR FILING DATE: 2000-02-25
      PRIOR APPLICATION NUMBER: 09/516,745
      PRIOR FILING DATE: 2000-03-01
      PRIOR APPLICATION NUMBER: 09/572,002
      PRIOR FILING DATE: 2000-05-14
      PRIOR APPLICATION NUMBER: 09/597,993
      PRIOR FILING DATE: 2000-06-19
      PRIOR APPLICATION NUMBER: 09/599,596
      PRIOR FILING DATE: 2000-06-22
      PRIOR APPLICATION NUMBER: 09/630,334
      PRIOR FILING DATE: 2000-07-31
      PRIOR APPLICATION NUMBER: 09/606,565
      PRIOR FILING DATE: 2000-06-29
      PRIOR APPLICATION NUMBER: 09/606,317
      PRIOR FILING DATE: 2000-06-29
      PRIOR APPLICATION NUMBER: 09/665,666
      PRIOR FILING DATE: 2000-09-20
      PRIOR APPLICATION NUMBER: 09/677,751
      PRIOR FILING DATE: 2000-09-30
      NUMBER OF SEQ ID NOS: 162
      SEQ ID NO 12
      LENGTH: 333
      TYPE: PRT
      ORGANISM: Homo sapiens
US-09-796-753-12

Query Match      100.0%; Score 943; DB 10; Length 333;
Best Local Similarity 100.0%; Pred. No. 1,6e-90;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GFDLYFLDKSGSVLHMHNEIYFVEQLAHKFTSPQLRMSFIVFSTRTTLMKLTEDREQ 60
      42 GFDLYFLDKSGSVLHMHNEIYFVEQLAHKFTSPQLRMSFIVFSTRTTLMKLTEDREQ 101
Db      61 IRGLEELQVLPFGDTYMHGFEFRASEQIYYENRQGYRRASVTIALTDGLHEDLFFYS 120
      102 IRGLEELQVLPFGDTYMHGFEFRASEQIYYENRQGYRRASVTIALTDGLHEDLFFYS 161
Qy      121 EREANRSDIGAIYCVGVDPFNETOLARIADSKDHVPVNDGFQALOGIIHSLKKSCT 180
      162 EREANRSDIGAIYCVGVDPFNETOLARIADSKDHVPVNDGFQALOGIIHSLKKSCT 221
Db      181 E 181
      222 E 222

RESULT 4
US-10-038-307-2
; Sequence 2, Application US/10038307
; Publication No. US20030134786A1
; GENERAL INFORMATION:
; APPLICANT: James B. ROTTMAN
```

APPLICANT: Theresa L. O'KEEFE
APPLICANT: Engin OZKAYNAK
APPLICANT: Judith J. HEALEY
TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
FILE REFERENCE: 7853-253-999
CURRENT APPLICATION NUMBER: US/10/038.307
CURRENT FILING DATE: 2002-06-28
NUMBER OF SEQ ID NOS: 26
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 333
TYPE: PRT
ORGANISM: Homo sapiens
US-10-038-307-2

Query Match 100.0%; Score 943; DB 14; Length 333;
Best Local Similarity 100.0%; Pred. No. 1.6e-90;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFDLYFLDKSGSVLHHMNEIYFVEQLAHKFI SPOLRMSFIVSTRGTTLMKLTEDREQ 60
DB 42 GFDLYFLDKSGSVLHHMNEIYFVEQLAHKFI SPOLRMSFIVSTRGTTLMKLTEDREQ 101
QY 61 IROGLEBLQVLPFGDITYMEHGEFRASEQIYENROGRTASVIALTDGLHEDLFFYS 120
DB 102 IROGLEBLQVLPFGDITYMEHGEFRASEQIYENROGRTASVIALTDGLHEDLFFYS 161
QY 121 EREANRSRDIGAIYVCVGDVFNFTOLARIADSKDHVPVNDGFOALOGIHSILKSCI 180
DB 162 EREANRSRDIGAIYVCVGDVFNFTOLARIADSKDHVPVNDGFOALOGIHSILKSCI 221
QY 181 E 181
DB 222 E 222

RESULT 5
US-10-201-292-2
Sequence 2, Application US/10201292
Publication No. US20030144193A1
GENERAL INFORMATION:
APPLICANT: James B. ROTTMAN
APPLICANT: Theresa L. O'KEEFE
APPLICANT: Engin OZKAYNAK
APPLICANT: Judith J. HEALEY
TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
FILE REFERENCE: 7853-253-999
CURRENT APPLICATION NUMBER: US/10/201.292
CURRENT FILING DATE: 2003-02-14
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 333
TYPE: PRT
ORGANISM: Homo sapiens
US-10-201-292-2

Query Match 100.0%; Score 943; DB 14; Length 333;
Best Local Similarity 100.0%; Pred. No. 1.6e-90;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFDLYFLDKSGSVLHHMNEIYFVEQLAHKFI SPOLRMSFIVSTRGTTLMKLTEDREQ 60
DB 42 GFDLYFLDKSGSVLHHMNEIYFVEQLAHKFI SPOLRMSFIVSTRGTTLMKLTEDREQ 101
QY 61 IROGLEBLQVLPFGDITYMEHGEFRASEQIYENROGRTASVIALTDGLHEDLFFYS 120
DB 102 IROGLEBLQVLPFGDITYMEHGEFRASEQIYENROGRTASVIALTDGLHEDLFFYS 161
QY 121 EREANRSRDIGAIYVCVGDVFNFTOLARIADSKDHVPVNDGFOALOGIHSILKSCI 180
DB 162 EREANRSRDIGAIYVCVGDVFNFTOLARIADSKDHVPVNDGFOALOGIHSILKSCI 221

QY 181 E 181
DB 222 E 222

RESULT 6
US-10-038-307-22
Sequence 22, Application US/10038307
Publication No. US20030134786A1
GENERAL INFORMATION:
APPLICANT: James B. ROTTMAN
APPLICANT: Theresa L. O'KEEFE
APPLICANT: Engin OZKAYNAK
APPLICANT: Judith J. HEALEY
TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
FILE REFERENCE: 7853-253-999
CURRENT APPLICATION NUMBER: US/10/038.307
CURRENT FILING DATE: 2002-06-28
NUMBER OF SEQ ID NOS: 26
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 22
LENGTH: 342
TYPE: PRT
ORGANISM: Homo sapiens
US-10-038-307-22

Query Match 100.0%; Score 943; DB 14; Length 342;
Best Local Similarity 100.0%; Pred. No. 1.7e-90;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFDLYFLDKSGSVLHHMNEIYFVEQLAHKFI SPOLRMSFIVSTRGTTLMKLTEDREQ 60
DB 51 GFDLYFLDKSGSVLHHMNEIYFVEQLAHKFI SPOLRMSFIVSTRGTTLMKLTEDREQ 110
QY 61 IROGLEBLQVLPFGDITYMEHGEFRASEQIYENROGRTASVIALTDGLHEDLFFYS 120
DB 111 IROGLEBLQVLPFGDITYMEHGEFRASEQIYENROGRTASVIALTDGLHEDLFFYS 170
QY 121 EREANRSRDIGAIYVCVGDVFNFTOLARIADSKDHVPVNDGFOALOGIHSILKSCI 180
DB 171 EREANRSRDIGAIYVCVGDVFNFTOLARIADSKDHVPVNDGFOALOGIHSILKSCI 230
QY 181 E 181
DB 231 E 231

RESULT 7
US-10-201-292-22
Sequence 22, Application US/10201292
Publication No. US20030144193A1
GENERAL INFORMATION:
APPLICANT: James B. ROTTMAN
APPLICANT: Theresa L. O'KEEFE
APPLICANT: Engin OZKAYNAK
APPLICANT: Judith J. HEALEY
TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
FILE REFERENCE: 7853-253-999
CURRENT APPLICATION NUMBER: US/10/201.292
CURRENT FILING DATE: 2003-02-14
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 22
LENGTH: 342
TYPE: PRT
ORGANISM: Homo sapiens
US-10-201-292-22

Query Match 100.0%; Score 943; DB 14; Length 342;
Best Local Similarity 100.0%; Pred. No. 1.7e-90;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFDLYFLDKSGSVLHHMNEIYFVEQLAHKFI SPOLRMSFIVSTRGTTLMKLTEDREQ 60

```

; SEQ ID NO 24
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-201-292-24
Query Match      100.0%; Score 943; DB 14; Length 345;
Best Local Similarity 100.0%; Pred. No. 1.7e-90;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      51 GFDLYFLDKSGSVLHHMNEIYFVEQLAHKFI SPOLRMSFIYFSTRGTTLMKLTEDREQ 110
        61 IROGLEELQKVLPGGDTYMHGFEFRASEQIYYENRQGYRTASVITALTDELHEDLFFYS 120
        111 IROGLEELQKVLPGGDTYMHGFEFRASEQIYYENRQGYRTASVITALTDELHEDLFFYS 170
Qy      121 EREANRSRDIGAIYVCVGVDFNETQIARLADSKDHVPVNDGFQALQGIHSLKKSCI 180
        171 EREANRSRDIGAIYVCVGVDFNETQIARLADSKDHVPVNDGFQALQGIHSLKKSCI 230
Db      181 E 181
        231 E 231

RESULT 8
US-10-038-307-24
; Sequence 24, Application US/10038307
; Publication No. US20030134786A1
; GENERAL INFORMATION:
; APPLICANT: James B. ROTTMAN
; APPLICANT: Theresa L. O'KEEFE
; APPLICANT: Engin OZKAYNAK
; APPLICANT: Judith J. HEALEY
; TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
; FILE REFERENCE: 7853-253-999
; CURRENT APPLICATION NUMBER: US/10/038.307
; CURRENT FILING DATE: 2002-06-28
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-038-307-24
Query Match      100.0%; Score 943; DB 14; Length 345;
Best Local Similarity 100.0%; Pred. No. 1.7e-90;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GFDLYFLDKSGSVLHHMNEIYFVEQLAHKFI SPOLRMSFIYFSTRGTTLMKLTEDREQ 60
        42 GFDLYFLDKSGSVLHHMNEIYFVEQLAHKFI SPOLRMSFIYFSTRGTTLMKLTEDREQ 101
        102 IROGLEELQKVLPGGDTYMHGFEFRASEQIYYENRQGYRTASVITALTDELHEDLFFYS 161
Qy      61 IROGLEELQKVLPGGDTYMHGFEFRASEQIYYENRQGYRTASVITALTDELHEDLFFYS 120
        102 IROGLEELQKVLPGGDTYMHGFEFRASEQIYYENRQGYRTASVITALTDELHEDLFFYS 161
Db      121 EREANRSRDIGAIYVCVGVDFNETQIARLADSKDHVPVNDGFQALQGIHSLKKSCI 180
        162 EREANRSRDIGAIYVCVGVDFNETQIARLADSKDHVPVNDGFQALQGIHSLKKSCI 221
Qy      181 E 181
        222 E 222

RESULT 9
US-10-201-292-24
; Sequence 24, Application US/10201292
; Publication No. US2003014193A1
; GENERAL INFORMATION:
; APPLICANT: James B. ROTTMAN
; APPLICANT: Theresa L. O'KEEFE
; APPLICANT: Engin OZKAYNAK
; APPLICANT: Judith J. HEALEY
; TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
; FILE REFERENCE: 7853-253-999
; CURRENT APPLICATION NUMBER: US/10/201.292
; CURRENT FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 24
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-201-292-24
Query Match      100.0%; Score 943; DB 14; Length 345;
Best Local Similarity 100.0%; Pred. No. 1.7e-90;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GFDLYFLDKSGSVLHHMNEIYFVEQLAHKFI SPOLRMSFIYFSTRGTTLMKLTEDREQ 60
        42 GFDLYFLDKSGSVLHHMNEIYFVEQLAHKFI SPOLRMSFIYFSTRGTTLMKLTEDREQ 101
        102 IROGLEELQKVLPGGDTYMHGFEFRASEQIYYENRQGYRTASVITALTDELHEDLFFYS 161
Qy      61 IROGLEELQKVLPGGDTYMHGFEFRASEQIYYENRQGYRTASVITALTDELHEDLFFYS 120
        102 IROGLEELQKVLPGGDTYMHGFEFRASEQIYYENRQGYRTASVITALTDELHEDLFFYS 161
Db      121 EREANRSRDIGAIYVCVGVDFNETQIARLADSKDHVPVNDGFQALQGIHSLKKSCI 180
        162 EREANRSRDIGAIYVCVGVDFNETQIARLADSKDHVPVNDGFQALQGIHSLKKSCI 221
Qy      181 E 181
        222 E 222

RESULT 10
US-09-833-245-621
; Sequence 621, Application US/09833245
; Publication No. US20040010134A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833.245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 621
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-245-621
Query Match      100.0%; Score 943; DB 11; Length 403;
Best Local Similarity 100.0%; Pred. No. 2.1e-90;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GFDLYFLDKSGSVLHHMNEIYFVEQLAHKFI SPOLRMSFIYFSTRGTTLMKLTEDREQ 60
        42 GFDLYFLDKSGSVLHHMNEIYFVEQLAHKFI SPOLRMSFIYFSTRGTTLMKLTEDREQ 101
        61 IROGLEELQKVLPGGDTYMHGFEFRASEQIYYENRQGYRTASVITALTDELHEDLFFYS 120
        102 IROGLEELQKVLPGGDTYMHGFEFRASEQIYYENRQGYRTASVITALTDELHEDLFFYS 161
Qy      121 EREANRSRDIGAIYVCVGVDFNETQIARLADSKDHVPVNDGFQALQGIHSLKKSCI 180
        162 EREANRSRDIGAIYVCVGVDFNETQIARLADSKDHVPVNDGFQALQGIHSLKKSCI 221
Db      181 E 181
        222 E 222

RESULT 11
```

US-10-201-292-28
/ Sequence 28, Application US/10201292
/ Publication No. US20030144193A1
/ GENERAL INFORMATION:
/ APPLICANT: James B. ROTTMAN
/ APPLICANT: Theresa L. O'KEEFE
/ APPLICANT: Engin OZKAYNAK
/ APPLICANT: Judith J. HEALEY
/ TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
/ FILE REFERENCE: 7853-253-999
/ CURRENT APPLICATION NUMBER: US/10/201,292
/ CURRENT FILING DATE: 2003-02-14
/ NUMBER OF SEQ ID NOS: 36
/ SOFTWARE: FaSeq for Windows Version 4.0
/ SEQ ID NO 28
/ LENGTH: 460
/ TYPE: PRN
/ ORGANISM: Homo sapiens
US-10-201-292-28

Query Match 100.0%; Score 943; DB 14; Length 460;
Best Local Similarity 100.0%; Pred. No. 2.5e-90;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFDLYFLDKSGSVLHMHNEIYFVEQLAHKFISPOLRMSFIVSTRGTTLMKLTEDREQ 60
DB 42 GFDLYFLDKSGSVLHMHNEIYFVEQLAHKFISPOLRMSFIVSTRGTTLMKLTEDREQ 101
QY 61 IROGLEBLOKVLPGGDTYMHGFERASEQIYENRQGRASVIALTDGELHEDLFFYS 120
DB 102 IROGLEBLOKVLPGGDTYMHGFERASEQIYENRQGRASVIALTDGELHEDLFFYS 161
QY 121 EREANRSRDLGAIYCVGVDPNETOLARIADSKDHYFPVNDGFOALOGIHSILKKS 180
DB 162 EREANRSRDLGAIYCVGVDPNETOLARIADSKDHYFPVNDGFOALOGIHSILKKS 221
QY 181 E 181
DB 222 E 222

RESULT 12
US-10-201-292-32
/ Sequence 32, Application US/10201292
/ Publication No. US20030144193A1
/ GENERAL INFORMATION:
/ APPLICANT: James B. ROTTMAN
/ APPLICANT: Theresa L. O'KEEFE
/ APPLICANT: Engin OZKAYNAK
/ APPLICANT: Judith J. HEALEY
/ TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
/ FILE REFERENCE: 7853-253-999
/ CURRENT APPLICATION NUMBER: US/10/201,292
/ CURRENT FILING DATE: 2003-02-14
/ NUMBER OF SEQ ID NOS: 36
/ SOFTWARE: FaSeq for Windows Version 4.0
/ SEQ ID NO 32
/ LENGTH: 479
/ TYPE: PRN
/ ORGANISM: Homo sapiens
US-10-201-292-32

Query Match 100.0%; Score 943; DB 14; Length 479;
Best Local Similarity 100.0%; Pred. No. 2.7e-90;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFDLYFLDKSGSVLHMHNEIYFVEQLAHKFISPOLRMSFIVSTRGTTLMKLTEDREQ 60
DB 42 GFDLYFLDKSGSVLHMHNEIYFVEQLAHKFISPOLRMSFIVSTRGTTLMKLTEDREQ 101
QY 61 IROGLEBLOKVLPGGDTYMHGFERASEQIYENRQGRASVIALTDGELHEDLFFYS 120
DB 102 IROGLEBLOKVLPGGDTYMHGFERASEQIYENRQGRASVIALTDGELHEDLFFYS 161

QY 121 EREANRSRDLGAIYCVGVDPNETOLARIADSKDHYFPVNDGFOALOGIHSILKKS 180
DB 162 EREANRSRDLGAIYCVGVDPNETOLARIADSKDHYFPVNDGFOALOGIHSILKKS 221
QY 181 E 181
DB 222 E 222

RESULT 13
US-10-201-292-34
/ Sequence 34, Application US/10201292
/ Publication No. US20030144193A1
/ GENERAL INFORMATION:
/ APPLICANT: James B. ROTTMAN
/ APPLICANT: Theresa L. O'KEEFE
/ APPLICANT: Engin OZKAYNAK
/ APPLICANT: Judith J. HEALEY
/ TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
/ FILE REFERENCE: 7853-253-999
/ CURRENT APPLICATION NUMBER: US/10/201,292
/ CURRENT FILING DATE: 2003-02-14
/ NUMBER OF SEQ ID NOS: 36
/ SOFTWARE: FaSeq for Windows Version 4.0
/ SEQ ID NO 34
/ LENGTH: 504
/ TYPE: PRN
/ ORGANISM: Homo sapiens
US-10-201-292-34

Query Match 100.0%; Score 943; DB 14; Length 504;
Best Local Similarity 100.0%; Pred. No. 2.9e-90;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFDLYFLDKSGSVLHMHNEIYFVEQLAHKFISPOLRMSFIVSTRGTTLMKLTEDREQ 60
DB 42 GFDLYFLDKSGSVLHMHNEIYFVEQLAHKFISPOLRMSFIVSTRGTTLMKLTEDREQ 101
QY 61 IROGLEBLOKVLPGGDTYMHGFERASEQIYENRQGRASVIALTDGELHEDLFFYS 120
DB 102 IROGLEBLOKVLPGGDTYMHGFERASEQIYENRQGRASVIALTDGELHEDLFFYS 161
QY 121 EREANRSRDLGAIYCVGVDPNETOLARIADSKDHYFPVNDGFOALOGIHSILKKS 180
DB 162 EREANRSRDLGAIYCVGVDPNETOLARIADSKDHYFPVNDGFOALOGIHSILKKS 221
QY 181 E 181
DB 222 E 222

RESULT 14
US-10-201-292-36
/ Sequence 36, Application US/10201292
/ Publication No. US20030144193A1
/ GENERAL INFORMATION:
/ APPLICANT: James B. ROTTMAN
/ APPLICANT: Theresa L. O'KEEFE
/ APPLICANT: Engin OZKAYNAK
/ APPLICANT: Judith J. HEALEY
/ TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
/ FILE REFERENCE: 7853-253-999
/ CURRENT APPLICATION NUMBER: US/10/201,292
/ CURRENT FILING DATE: 2003-02-14
/ NUMBER OF SEQ ID NOS: 36
/ SOFTWARE: FaSeq for Windows Version 4.0
/ SEQ ID NO 36
/ LENGTH: 529
/ TYPE: PRN
/ ORGANISM: Homo sapiens
US-10-201-292-36

```
Query Match      100.0%; Score 943; DB 14; Length 529;
Best Local Similarity 100.0%; Pred. No. 3.3e-90;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GFDLYFLIDKSGSVLHMHNEIYFVEQLAHKFISPOLRMSFIVSTRGTTLMKLTEDREQ 60
      |||
Db      42 GFDLYFLIDKSGSVLHMHNEIYFVEQLAHKFISPOLRMSFIVSTRGTTLMKLTEDREQ 101
      |||

Qy      61 IRQGLEELQVLPQGDITYMEHGFEFRASEQIYENRQGYRTASVIALTDGELHEDLFFYS 120
      |||
Db      102 IRQGLEELQVLPQGDITYMEHGFEFRASEQIYENRQGYRTASVIALTDGELHEDLFFYS 161
      |||

Qy      121 EREANRSRDIGAIYVCVGDVDFNETOLARIADSKDHVPVNDGFOALQGIHSLKKSCT 180
      |||
Db      162 EREANRSRDIGAIYVCVGDVDFNETOLARIADSKDHVPVNDGFOALQGIHSLKKSCT 221
      |||

Qy      181 E 181
      |
Db      222 E 222

RESULT 15
US-10-038-307-18
; Sequence 18; Application US/10038307
; Publication No. US20030134786A1
; GENERAL INFORMATION:
; APPLICANT: James B. ROTTMAN
; APPLICANT: Theresa L. O'KEEFE
; APPLICANT: Engin OZKAYNAK
; APPLICANT: Judith J. HEALEY
; TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
; FILE REFERENCE: 7853-253-999
; CURRENT APPLICATION NUMBER: US/10/038.307
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 551
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-038-307-18

Query Match      100.0%; Score 943; DB 14; Length 551;
Best Local Similarity 100.0%; Pred. No. 3.3e-90;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GFDLYFLIDKSGSVLHMHNEIYFVEQLAHKFISPOLRMSFIVSTRGTTLMKLTEDREQ 60
      |||
Db      42 GFDLYFLIDKSGSVLHMHNEIYFVEQLAHKFISPOLRMSFIVSTRGTTLMKLTEDREQ 101
      |||

Qy      61 IRQGLEELQVLPQGDITYMEHGFEFRASEQIYENRQGYRTASVIALTDGELHEDLFFYS 120
      |||
Db      102 IRQGLEELQVLPQGDITYMEHGFEFRASEQIYENRQGYRTASVIALTDGELHEDLFFYS 161
      |||

Qy      121 EREANRSRDIGAIYVCVGDVDFNETOLARIADSKDHVPVNDGFOALQGIHSLKKSCT 180
      |||
Db      162 EREANRSRDIGAIYVCVGDVDFNETOLARIADSKDHVPVNDGFOALQGIHSLKKSCT 221
      |||

Qy      181 E 181
      |
Db      222 E 222

Search completed: June 13, 2005, 20:36:52
Job time : 59.0354 secs
```


B.

A/Reference number: A31108; MUID:88315033; PMID:2457584

A/Accession: A31108

A/Molecule type: mRNA

A/Residues: 1-1153 <COR>

A/Cross-references: UNIPROT:P11215; GB:J03925; NID:g187284; P1DN:AAA59544.1; P1D:g3071448

A/Note: part of this sequence was confirmed by protein sequencing

A/Arnaut, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.

J. Cell Biol. 106, 2153-2158, 1988

A/Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor Mac-1 (complement receptor type 3, CD11b).

A/Reference number: A28915; MUID:88257215; PMID:2454931

A/Accession: A28915

A/Molecule type: mRNA

A/Residues: 1-499, 501-965, 'P', 967-1153 <ARN>

A/Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:g186935; P1DN:AAA594

A/Note: the authors translated the codon TAC for residue 1129 as Thr

A/Note: part of this sequence, including the amino end of the mature protein, was confirmed

A/Shelly, C.S.; Arnaut, M.A.

Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991

A/Title: The promoter of the CD11b gene directs myeloid-specific and developmentally regulated

A/Reference number: A41600; MUID:92073318; PMID:1683702

A/Accession: A41600

A/Molecule type: DNA

A/Residues: 1-9 <SHE>

A/Cross-references: GB:M76724; NID:g180018; P1DN:AAA8410.1; P1D:g553215

A/Arnaut, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.

Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988

A/Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhesion

A/Reference number: A94193; MUID:88190151; PMID:2833753

A/Accession: A30892

A/Molecule type: mRNA

A/Residues: 917-1042 <AR2>

A/Cross-references: GB:M18044

A/Hickstein, D.D.; Hickey, M.J.; Ocole, J.; Baker, D.M.; Back, A.L.; Roth, G.J.

R. Hickett, A.L.; Hickey, M.J.; Ocole, J.; Baker, D.M.; Back, A.L.; Roth, G.J.

Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989

A/Title: cDNA sequence for the alpha subunit of the human neutrophil adherence receptor

A/Reference number: A32218; MUID:89098893; PMID:2563162

A/Accession: A32218

A/Molecule type: mRNA

A/Residues: 9-1153 <HIC>

A/Cross-references: GB:J04145; NID:g189068; P1DN:AAA59903.1; P1D:g386975

A/Note: part of this sequence was confirmed by protein sequencing

A/Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.

J. Immunol. 150, 480-490, 1993

A/Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-in

during evolution.

A/Reference number: A46526; MUID:93123748; PMID:8419480

A/Accession: A46526

A/Status: not compared with conceptual translation

A/Molecule type: DNA

A/Residues: 1-499, 501-1153 <FLB>

A/Cross-references: GB:S52227; NID:g263047; P1DN:AA24821.1; P1D:g263049

A/Note: the last three bases of intron 13, CAG, are included in some but not all mature

A/Note: sequence extracted from NCBI backbone (NCBI:121963)

A/Pierce, M.W.; Remold-O'Donnell, E.; Todd III, R.F.; Arnaut, M.A.

Biochem. Biophys. Acta 874, 368-371, 1986

A/Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across sp

A/Reference number: A90664; MUID:87076671; PMID:3539202

A/Accession: A26091

A/Molecule type: protein

A/Residues: 17-31 <PIR>

A/Experimental source: granulocytes

R. Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G.

Blood 79, 865-870, 1992

A/Title: Characterization of the myeloid-specific CD11b promoter.

A/Reference number: 152567; MUID:92144986; PMID:1346576

A/Accession: 152567

A/Status: translated from GB/EMBL/DBU

A/Molecule type: DNA

A:Residues: 1-9 <RES>
A:Cross-references: GB: M64477; NID: g180184; PIDN: AAA51960.1; PID: g953219
C:Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1 c
C:Genetics:
A:Gene: GDB: ITGAM; CR3A
A:Cross-references: GDB: 120599; OMIM: 120980
A:Map position: 16p11.2-16p11.2
A:Note: promoter containing a GATA motif and two Sp1 consensus binding sites
C:Superfamily: cell surface glycoprotein CD18, von Willebrand factor type A repeat homo
C:Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; mag
F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-1153/Product: cell surface glycoprotein CD18 #status experimental <MAT>
F:148-318/Domain: extracellular #status predicted <EXT>
F:148-318/Domain: von Willebrand factor type A repeat homology <VMA2>
F:465-473/Region: calcium/magnesium binding #status predicted
F:530-538/Region: calcium/magnesium binding #status predicted
F:593-601/Region: calcium/magnesium binding #status predicted
F:1109-1134/Domain: transmembrane #status predicted <TM>
F:1135-1153/Domain: intracellular #status predicted <INT>
F:86,240,391,469,693,697,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Binding

Query Match 14.1%; Score 132.5; DB 1; Length 1153;
Best Local Similarity 26.3%; Pred. No. 0.0026;
Matches 52; Conservative 41; Mismatches 66; Indels 39; Gaps 11;

Qy 3 DLVPIIDKSGSVL-HHNNHYYFVQQLAKRFSPOLRMFSIVSTGKTIIMKTED---- 57
Db 150 DIAFLIDSGSITPPDFFRMKEFVST----VWEOLKSKETLPS----LMQSEERIH 199
Oy 58 -----REQIRQGLLEIQLVPGDDTYMEHGFEFASQIYYENRQGYRTA-SVIALT 108
Db 200 FTFKPEQNPNPRLSLVKPTQLL--GRHTTAQIRKRVARELPFITNGARCAAPFLIVIT 257
Qy 109 DGEIHEDLFYSE--RANRSRLDALVYCVGKDFENETQLA-----IADS--KDHVP 159
Db 258 DGEKRGDPLGVEDVPEADRE---GVIRYVIGVDAFRSEKSHQELNTITASKPRDHVFQ 314
Oy 160 VNDGFOALQGITHSILTK 177
Db 315 VNN-FEALKTIONQLREK 331

RESULT 3
A40020
collagen alpha 1(XII) chain precursor - chicken
M:Alternate names: fibrochimerin
C:Species: Gallus gallus (chicken)
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #ext change 09-Jul-2004
C:Accession: A40020, A34485, B34485, A28037, S23814, S22254, S28811
R:Yamagata, M.; Yamada, K.M.; Yamada, S.S.; Shinomura, T.; Tanaka, H.; Nishida, Y.; Obayashi, J.; Cell Biol. 115, 209-221, 1991
A:Title: The complete primary structure of type XII collagen shows a chimeric molecule w
ous region of type IX collagen, and short collagenous domains with an Arg-Gly-Asp site.
A:Reference number: A40020; MUID: 92011862; PMID: 1918137
A:Accession: A40020
A:Molecule type: mRNA
A:Residues: 1-3124 <YMW>
A:Cross-references: UNIPROT: P13944; GB: D00824; NID: g222810; PIDN: BAA00701.1; PID: g222811
A:Note: in the authors' translation residues 1216-1219 are shown after residue 1235 and,
R:Gordon, M.K.; Gerecke, D.R.; Dublet, B.; van der Rest, M.; Olsen, B.R.
J. Biol. Chem. 264, 19772-19778, 1989
A:Title: Type XII collagen. A large multidomain molecule with partial homology to type IX
A:Reference number: A34485; MUID: 90062079; PMID: 2584192
A:Accession: A34485
A:Molecule type: mRNA
A:Residues: 2456-2758, 'A', 2760-2802, 'F', 2804-2976, 'F', 2978-3124 <GOR>
A:Cross-references: EMBL: J05137; NID: g211284; PIDN: AAA48635.1; PID: g211285
A:Accession: B34485
A:Molecule type: protein
A:Residues: 2772-2792, 2846-2873 <GOR2>
R:Gordon, M.K.; Gerecke, D.R.; Olsen, B.R.
Proc. Natl. Acad. Sci. U.S.A. 84, 6040-6044, 1987
A:Title: Type XII collagen: distinct extracellular matrix component discovered by cDNA c

Matches 49; Conservative 42; Mismatches 79; Indels 16; Gaps 9;

Qy 3 DLYFIILDKSGSV-LHHNNEIYFVEQLAHK-F--ISPC-LMSFVFEVSTRGTTLMKLT-R-E 56
Db 2327 DLYFIILDKSGSV-LHHNNEIYFVEQLAHK-F--ISPC-LMSFVFEVSTRGTTLMKLT-R-E 56

Qy 57 DREQIROGELBELQKVLFGCDTYMHGEGFERASEQIY-YENROGYRTASVIALITDGEIHD 115
Db 2387 DREQIROGELBELQKVLFGCDTYMHGEGFERASEQIY-YENROGYRTASVIALITDGEIHD 115

Qy 116 LFFYSEREANRSRDGLAIYVCVGVKDFNETQIARIAD-SKDHVPVNDGFQALQGIHS 173
Db 2445 V---RKAATVYIHSQSGSVFVGVADVVDNELAKIASPESRHVFIYVD-FDAFEKIQDN 2499

Qy 174 ILKSC 179
Db 2500 LVPFVC 2505

RESULT 4
A45638
Immunodominant microneme protein Etp100 - Eimeria tenella
C/Spectes: Eimeria tenella
C/Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C/Accession: A45638
R:Tomley, F.M.; Clarke, L.E.; Kawazoe, U.; Djikema, R.; Kok, J.J.
Mol. Biochem. Parasitol. 49, 277-288, 1991
A>Title: Sequence of the gene encoding an immunodominant microneme protein of Eimeria t
A/Reference number: A45638; MUID:92131064; PMID:1175171
A/Accession: A45638
A>Status: preliminary
A/Molecule type: DNA
A/Residues: 1-712 <TOM>
A/Cross-references: UNIPROT:O43981; GB:AF032905; GB:M73495; NID:92707732; PIDN:AMD03350
A/Note: Sequence extracted from NCBI backbone (NCBIN:77752, NCBI:P:77756)
F:48-218/Domain: von Willebrand factor type A repeat homology <VWA1>
F:238-286/Domain: thrombospondin type 1 repeat homology <THR1>
F:309-371/Domain: thrombospondin type 1 repeat homology <THR2>
F:372-432/Domain: thrombospondin type 1 repeat homology <THR3>
F:433-453/Domain: thrombospondin type 1 repeat homology <THR4>
F:494-556/Domain: thrombospondin type 1 repeat homology <THR5>
F:560-610/Domain: thrombospondin type 1 repeat homology <THR6>

Query Match 13.5%; Score 127.5; DB 2; Length 712;
Best Local Similarity 23.7%; Pred. No. 0.0039;
Matches 47; Conservative 40; Mismatches 76; Indels 35; Gaps 10;

Qy 3 DLYFIILDKSGSV-LHHNNEIYFVEQLAHK-F--ISPC-LMSFVFEVSTRGTTLMKLTEDRE 59
Db 50 DLYFIILDKSGSV-LHHNNEIYFVEQLAHK-F--ISPC-LMSFVFEVSTRGTTLMKLTEDRE 59

Qy 60 QIROGELBELQKVLFGCDTYMHGEGFERASEQIY-YENROGYRTASVIALITDGEIHD 104
Db 99 KVRNKLNDPKATNSLISASRLSYSTGVLYTHYGLQDA-KLLLYTNAGARANNVPRLV 157

Qy 105 IALTDEIHEHDLFFYSEREANRSRDGLAIYVCVGV-KDFNETQIARIADSKDHVP-VN 161
Db 158 LMTDGA--SVLPQRTSSAALARDAGAVLVGLGVSSSECRSIAGCSTNCPRYLQ 215

Qy 162 DGFQALQGIHSILKSC 179
Db 216 SNMSNVTVQVNGIITKAC 233

RESULT 5
RKHUIC
cell surface glycoprotein CD1c precursor - human
N/Alternate names: leukocyte adhesion receptor p150,95 alpha chain
C/Spectes: Homo sapiens (man)
C/Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C/Accession: A3584; A3543; S00864
R:Corbdi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 12750-12751, 1990

Oy 105 IALTDLGHELDLFFYSREANRSRDLCALIVCVGVKQ-FNETQLARIADSKDHVPVNDG 163
Db 624 IITLTDSQDNV----TGPAIDSAKSLISINTFAIGVTDHVLASLESIASFPNRFVY-DK 678
Oy 164 FOALOGIHSILKKS 178
Db 679 FKOLDTRLRSMIQA 693

RESULT 8

C2MS

Classical-complement-pathway C3/C5 convertase (EC 3.4.21.43) C2 component precursor - mcd
N/Alternate names: C3 convertase; C5 convertase; complement C2

C/Species: Mus musculus (house mouse)

C/Date: 30-Sep-1993 #sequence revision 30-Sep-1993 #text_change 09-Jul-2004

C/Accession: A38876; B36593; I54429

R/Ishtawa, N.; Nonaka, M.; Wetzel, R.A.; Colten, H.R.

submitted to GenBank, January 1991

A/Reference number: A38875

A/Accession: A38876

A/Molecule type: DNA

A/Residues: 1-760 <IS2>

A/Cross-references: UNIPROT:P21180; GB:M57891; GB:J05661; NID:g192436; PIDN:AAA63294.1;

J. Biol. Chem. 265, 19040-19046, 1990

A/Title: Murine complement C2 and factor B genomic and cDNA cloning reveals different me

A/Reference number: A36593; MUID:91035430; PMID:2229060

A/Accession: B36593

A/Molecule type: mRNA

A/Residues: 1-760 <ISH>

A/Cross-references: EMBL:M57891; NID:g192436; PIDN:AAA63294.1; PID:g192437

R/Ishtawa, N.; Wetzel, R.A.; Colten, H.R.

Immunogenetics 25, 290-298, 1987

A/Title: DNA polymorphism of H2C III genes in inbred and wild mouse strains.

A/Reference number: I54429; MUID:87192938; PMID:2883115

A/Accession: I54429

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 660-677, 'R', 679, 681-723, 'G', 725 <RES>

A/Cross-references: GB:M16271; NID:g199289; PIDN:AAA9562.1; PID:g199290

C/Genetics:

A/Introns: 16/1, 91/1, 153/1, 212/1, 245/3, 290/3, 337/1, 384/1, 414/1, 461/1, 492/3, 51

C/Complex: The proenzyme forms a complex with C4a and is activated by cleavage into C2a,

C/Function:

A/Description: cleaves complement C3 and complement C5 alpha chains

C/Suprafamily: complement B/C2; complement factor H repeat homology; tryptan homology; v

F/1-18/DNA: signal sequence; complement predicted <SIG>

F/19-250/Product: complement C2b fragment #status predicted <C2b>

F/22-69/DNA: complement factor H repeat homology <FH1>

F/94-149/DNA: complement factor H repeat homology <FH2>

F/156-210/DNA: complement factor H repeat homology <FH3>

F/551-160/Product: complement C2a fragment long form #status predicted <C2a>

F/551-605, 613-760/Product: complement C2a fragment short form #status predicted <C2s>

F/559-449/DNA: von Willebrand factor type A repeat homology <VFA>

F/478-747/DNA: tryptan homology #status atypical <TRY>

F/22-69, 48-89, 94-136, 122-149, 156-197, 182-210, 470-550, 499-515, 593-609, 647-674, 685-715/DNA

F/27, 117, 297, 340, 474, 478, 663/Binding site: carbohydrate (Asn) (covalent) #status predict

F/514, 570, 689/Active site: His, Asp, Ser #status predicted

Query Match 12.1%; Score 114.5; DB 1; Length 760;

Best Local Similarity 24.3%; Pred. No. 0.056; Mismatches 72; Indels 49; Gaps 10;

Matches 52; Conservative 41; Mismatches 72; Indels 49; Gaps 10;

Oy 3 DLVFIIDKSGSVLHWNIEY-YVEEQALAHKFIISPOLR-MSFVIFSTRGTLMKLTEDRE 59

Db 261 NLVLLDASQSVTEKFDIFKSAELMVERIPSEFVAVVAILTFASQPKTMSILSERS 320

Oy 60 QINQGLBELOKVLPGDPTVHNEGFERSAQIYVE-----NRGYSRTAS---- 102

Db 321 Q-----DVEVITSLDSASVSKDHENATYEVLLIRVYSMOTOMDRMGMTSAWKEI 374

Oy 103 --VIALTDGLHEDLFFYSREANRSRDLCAL-----VYCVGV-----KQFNE 144
Db 375 RHITLLTDGK--SNMGDSPKKAVTRIRBELISIQNRNDYDIYAIGVKLDVWKEINE 432
Oy 145 TQLARIADSKDHVPVNDGFOALOGIHSILKKS 178
Db 433 --LQSKKQGERHAFILQDA-KALQQLFEHMLDV 463

RESULT 9

A45974

collagen alpha 1(XIV) chain precursor, short form 2 - chicken

N/Alternate names: undulin

C/Species: Gallus gallus (chicken)

C/Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 09-Jul-2004

C/Accession: A45974; S30085; S22916; S17035; S20833

R/Gerecke, D.R.; Foley, J.W.; Castagnola, P.; Gennari, M.; Dublet, B.; Cancedda, R.; Lin

J. Biol. Chem. 268, 12177-12184, 1993

A/Title: Type XIV collagen is encoded by alternative transcripts with distinct 5' region

ns.

A/Reference number: A45974; MUID:93280195; PMID:8505337

A/Accession: A45974

A/Status: preliminary

A/Molecule type: mRNA; protein

A/Residues: 1-1747 <GER>

A/Cross-references: UNIPROT:P32018

A/Experimental source: embryo skin

A/Note: sequence inconsistent with the nucleotide translation

R/Apple, S.S.

submitted to the EMBL Data Library, March 1992

A/Reference number: S30085

A/Accession: S30085

A/Molecule type: mRNA

A/Residues: 1472-1660 <APT>

A/Cross-references: EMBL:X65122; NID:962871; PIDN:CAA46238.1; PID:g938175

R/Trueb, J.; Trueb, B.

Eur. J. Biochem. 207, 549-557, 1992

A/Title: Type XIV collagen is a variant of undulin.

A/Reference number: S22916; MUID:92339443; PMID:1339349

A/Accession: S22916

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 286-494, 'Q', 496-834, 'A', 836-1119, 'KL', 1122-1402, 1409-1439 <TRU>

R/Gordon, M.K.; Castagnola, P.; Dublet, B.; Linsemayer, T.F.; van der Rest, M.; Mayne, F

Eur. J. Biochem. 201, 333-338, 1991

A/Title: Cloning of a cDNA for a new member of the class of fibril-associated collagens

A/Reference number: S17035; MUID:92037585; PMID:1935930

A/Accession: S17035

A/Molecule type: mRNA

A/Residues: 1472-1659 <GOR1>

A/Accession: S20833

A/Molecule type: protein

A/Residues: 1551-1570, 1593-1599, 1639-1667 <GOR2>

C/Keywords: alternative splicing; coiled coil; extracellular matrix; glycoprotein; trimer

F/40-204/DNA: von Willebrand factor type A repeat homology <VWA1>

F/236-409/DNA: fibronectin type III repeat homology <FN3A>

F/336-409/DNA: fibronectin type III repeat homology <FN3B>

F/418-498/DNA: fibronectin type III repeat homology <FN3C>

F/507-591/DNA: fibronectin type III repeat homology <FN3D>

F/625-707/DNA: fibronectin type III repeat homology <FN3E>

F/716-798/DNA: fibronectin type III repeat homology <FN3F>

F/806-893/DNA: fibronectin type III repeat homology <FN3G>

F/924-1089/DNA: von Willebrand factor type A repeat homology <VWA2>

F/1111-1352/DNA: non-collagenous NC4 #status predicted <NC4>

F/1554-1659/DNA: triple helical domain COL1 #status predicted <COL1>

Query Match 12.0%; Score 113; DB 2; Length 1747;

Best Local Similarity 26.1%; Pred. No. 0.2; Mismatches 80; Indels 20; Gaps 11;

Matches 49; Conservative 39; Mismatches 80; Indels 20; Gaps 11;

Oy 3 DLVFIIDKSGSV-LHWNIEYVEEQAL-AHKFISPD-LRMSFVIFSTRGTLMKLT--E 56

Db 442 LASKUNQHVFVKD-MENLEDFYQWIDES 472

RESULT 15

821369

collagen alpha 2(VI) chain precursor - mouse

C/Species: Mus musculus (house mouse)

C/Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 09-Jul-2004

C/Accession: S21369; S28808; S13745; S18863

R/Author: X.Y.Z.; Bardou, X.Y.Z.; Dani, C.

A/Reference number: S21369

A/Accession: S21369

A/Molecule type: mRNA

A/Residues: 1-1029 <IBR>

A/Cross-references: UNIPROT:Q02788; EMBL:X65582; NID:G49808; PIDN:CAA46541.1; PID:G49805

R/Author: A.; Bertrand, B.; Bardou, S.; Amri, E.Z.; Grimaldi, P.; Allhaud, G.; Dani, C.

B/Author: J. 289, 141-147, 1993

A/Title: Cloning of alpha2 chain of type VI collagen and expression during mouse develop

A/Reference number: S28808; MUID:93143659; PMID:8380980

A/Accession: S28808

A/Molecule type: mRNA

A/Residues: 266-1029 <IB2>

A/Cross-references: EMBL:X62332; NID:G49906; PIDN:CAA44206.1; PID:G49907

R/Author: C.D.; Jimenez, S.A.

Macrix 11, 1-9, 1991

A/Title: Structure of cDNAs encoding the triple-helical domain of murine alpha-2(VI) col

rate oligonucleotides for generation of novel cDNA clones.

A/Reference number: S13745; MUID:9126374; PMID:1709252

A/Accession: S13745

A/Molecule type: mRNA

A/Residues: 266-267, 'S', 269-294, 'L', 296-600 <CON>

A/Cross-references: GB:U06343; NID:G192671; PIDN:AAA37441.1; PID:G192672

A/Note: the sequence from Fig. 3 is inconsistent with that from Fig. 4 in having 306-116

C/superfamily: collagen VI; von Willebrand factor type A repeat homology

C/keywords: alternative splicing; cartilage; cell binding; connective tissue; disulfide

F/1-28/Domain: signal sequence #status predicted <SIG>

F/29-1029/Product: collagen alpha 2(VI) chain #status predicted <MAT>

F/29-265/Domain: globular #status predicted <NC2>

F/54-226/Domain: von Willebrand factor type A repeat homology <VWA1>

F/266-600/Domain: collagenous #status predicted <COL>

F/376-378/Region: cell attachment (R-G-D) motif

F/436-438/Region: cell attachment (R-G-D) motif

F/499-501/Region: cell attachment (R-G-D) motif

F/508-510/Region: cell attachment (R-G-D) motif

F/549-551/Region: cell attachment (R-G-D) motif

F/601-1029/Domain: globular #status predicted <NC1>

F/623-799/Domain: von Willebrand factor type A repeat homology <VWA2>

F/841-1012/Domain: von Willebrand factor type A repeat homology <VWA3>

F/150,337,640,907,963/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 11.24; Score 105.5; DB 1; Length 1029;

Best Local Similarity 23.18; Pred. No. 0.48;

Matches 49; Conservative 41; Mismatches 85; Indels 37; Gaps 8;

```

QY 3 DLYFIDKSGSV-----LHHMNE-IYFVFOALHKFISPOLBMSF-----IVFSTRG 48
DB 56 NVYFVLDTSVAMQSTDLNMQQFPOFISQLQNEFFYLDQVALSMRYGGIHFSDQV 115
QY 49 TLMKLTEDREQROGLEBLQKVLPGSDTYMHGFEFASBOIYYENRQGYRTASVIALT 108
DB 116 EVSPSPGSDASFTKSLQGRSPFRG--TPTDCLAMMTQOIRGHVGKGVNFAVVI--T 171
QY 109 DGLHEDLPFYSERENRSDLAIVYCVGV-KDFNETOLARIADSKOHVFPVN----- 161
DB 172 DGHVTSFGCGIKQARAREEGIRLPALAPNRNLNEQGLRIANSFHELYRNRYATMRP 231
QY 162 DGFOLQGIHSHIL-----KKSCLIE 181
DB 232 DSTEIDDTINRIIKWKHEAYGECYKVCLE 263

```

Search completed: June 13, 2005, 20:06:38
Job time : 13.9858 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: June 13, 2005, 19:37:47 ; Search time 60.7608 Seconds
(without alignments)
1525.429 Million cell updates/sec

Title: US-09-970-076-2_COPY_42_222

Perfect score: 943
Sequence: 1 GFDLYRILDKSGSVLHWMNE.....DGFQALQGIHILKSKCIE 161

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: UniProt_03.*
1: unidprot_sprot.*
2: unidprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	943	100.0	564	1 ATR1_HUMAN	Q9h6x2 homo sapien
2	934	99.0	562	1 ATR1_MOUSE	Q9c252 mus musculu
3	542	57.5	489	1 ATR2_HUMAN	P58335 homo sapien
4	540	57.3	487	2 O6DFX2	O6dfx2 mus musculu
5	375	39.8	641	2 Q8BVM2	Q8bvm2 mus musculu
6	146	15.5	1332	2 Q9BPQ8	Q9bpq8 halocynthia
7	141.5	15.0	1161	1 ITAD_RAT	Q8qy67 rattus norv
8	139	14.7	441	2 O8T6U5	O8t6u5 mytilus edu
9	139	14.7	441	2 O8T6U5	O8t6u5 mytilus edu
10	136	14.4	453	2 O8T5C2	O8t5c2 mytilus gal
11	136	14.4	724	2 O045B8	O045b8 elmeria max
12	134.5	14.3	3119	1 CA1C_MOUSE	Q60847 mus musculu
13	132.5	14.1	1152	1 ITAM_HUMAN	P11215 homo sapien
14	129.5	13.7	1182	2 O8C6K9	O8c6k9 mus musculu
15	128	13.6	7124	1 CA1C_CHICK	P13944 gallus gall
16	127.5	13.5	7124	2 Q439B1	Q439b1 elmeria ten
17	127.5	13.5	765	2 Q9UB09	Q9ub09 neospora ca
18	126.5	13.4	3063	1 CA1C_HUMAN	Q99715 homo sapien
19	125.5	13.3	1163	1 ITAX_MOUSE	P20702 homo sapien
20	124.5	13.2	1153	1 ITAM_MOUSE	P05555 mus musculu
21	124.5	13.2	1162	1 ITAD_HUMAN	O13349 homo sapien
22	122	12.9	517	2 Q438E3	Q438e3 homo sapien
23	121.5	12.9	920	2 Q289E4	Q289e4 sus scrofa
24	119	12.6	1626	2 Q8NFW1	Q8nfw1 homo sapien
25	117.5	12.5	760	2 Q70350	Q70350 mus musculu
26	117	12.4	1823	2 Q7PRP5	Q7prp5 anopheles g
27	116.5	12.4	919	2 Q75R52	Q75r52 lymanaea ata
28	116	12.3	637	2 Q8IYX1	Q8iyx1 homo sapien
29	116	12.3	1284	2 Q6PI59	Q6pi59 homo sapien
30	116	12.3	1329	1 RFI0_HUMAN	Q9p218 homo sapien
31	116	12.3	2944	2 Q63870	Q63870 mus musculu

32	116	12.3	3183	2 Q65ZC2	Q65zc2 caenorhabdi
33	116	12.3	3767	1 MU03_CAEEL	P34576 caenorhabdi
34	114.5	12.1	760	1 CO2_MOUSE	P31180 mus musculu
35	114.5	12.1	2104	2 Q212B1	Q212b1 caenorhabdi
36	114.5	12.1	2104	2 Q964N4	Q964n4 caenorhabdi
37	113.5	12.0	790	2 Q6DC06	Q6dc06 xenopus lae
38	113	12.0	599	2 Q8WVQ1	Q8wvq1 bolitena vi
39	113	12.0	1888	1 CA1B_CHICK	P32018 gallus gall
40	112	11.9	588	2 Q6ZED8	Q6zed8 synechocyst
41	112	11.9	652	2 Q95LR2	Q95lr2 bos taurus
42	111.5	11.8	929	1 CA1C_NOTVI	Q91145 notophthalm
43	111	11.8	440	2 Q8CBT2	Q8cbt2 mus musculu
44	111	11.8	848	2 Q8C720	Q8c720 mus musculu
45	111	11.8	3567	2 Q9ES77	Q9es77 mus musculu

ALIGNMENTS

RESULT 1
ID ATR1_HUMAN STANDARD: PRT; 564 AA.
AC Q9h6x2; Q96P02; Q9NVP3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Anthrax toxin receptor 1 precursor (Tumor endothelial marker 8).
GN Name=ANTXR1; Synonyms=ATR, TEM8;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=20407466; PubMed=10947988; DOI=10.1126/science.289.5482.1197;
RA St Croix B., Rago C., Velculescu V., Traverso G., Romans K.B.,
RA Montgomery E., Lai A., Riggs G.J., Lengauer C., Vogelstein B.,
RA Kinzler K.W.;
RT "Genes expressed in human tumor endothelium";
RL Science 289:1197-1202(2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2), AND INTERACTION WITH ANTHRAX TOXIN.
RX MEDLINE=11557240; PubMed=11700562; DOI=10.1038/n510198;
RA Bradley K.A., Mogridge J., Mourez M., Collier R.J., Young J.A.T.;
RT "Identification of the cellular receptor for anthrax toxin";
RL Nature 414:225-229(2001).
RN [3]
RP SEQUENCE OF 184-564 FROM N.A. (ISOFORM 1), AND SEQUENCE FROM N.A.
(ISOFORM 3).
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Makamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto K., Ii S., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Nagahara K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Puruya T., Kikawa E.,
RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoaka S., Chiba Y.,
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yoshida T.,
RA Kuwano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Togawa S., Komai F., Haru R., Takeuchi K., Arita M.,
RA Imose N., Mueselino K., Yuki H., Oshima A., Saeki N., Aoeuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tanigami A., Tanigami A., Fujisawa T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiro M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,

RA Matsunura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano Y., Seton T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuo Y., Yamauchi R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.,
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RL Nat. Genet. 36:40-45(2004).
RN [4]
RN SEQUENCE FROM N.A. (ISOFORM 4).
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Adamson R.D., Mulhally S.J.,
RA Bosak S.A., McEwan P.U., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.U., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W.,
RA Butcherfield Y.S.N., Krzywinski M.I., Skalek A., Smillie D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RN INTERACTION WITH ANTHRAX TOXIN.
RC TISSUE=Placenta;
RX MEDLINE=22608610; PubMed=12700348; DOI=10.1073/pnas.0431098100;
RA Scobie H.M., Rainey G.J.A., Bradley K.A., Young J.A.T.,
RT "Human capillary morphogenesis protein 2 functions as an anthrax toxin
RT receptor.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:5170-5174(2003).
RN [6]
RN SLICE ISOFORM(S) THAT ARE POTENTIAL NMD TARGET(S).
RX PubMed=14759258; DOI=10.1186/gb-2004-5-2-r8;
RA Hillman R.T., Green R.E., Brenner S.E.,
RL "An unappreciated role for RNA surveillance.";
CC Genome Biol. 5:RESEARCH008.1-RESEARCH008.16(2004).
CC -1- FUNCTION: Cellular role is not yet known.
CC -1- SUBUNIT: Binds to the protective antigen (PA) of *Bacillus*
CC anthracis. Binding does not occur in the presence of calcium.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Probable).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoform=4;
CC Comment=Experimental confirmation may be lacking for some
CC isoforms;
CC Name=1;
CC IsoId=Q9H6X2-1; Sequence=Displayed;
CC Note=May be produced at very low levels due to a premature stop
CC codon in the mRNA, leading to nonsense-mediated mRNA decay;
CC Name=2;
CC IsoId=Q9H6X2-2; Sequence=VSP_000444; VSP_000445;
CC Name=3;
CC IsoId=Q9H6X2-3; Sequence=VSP_000446; VSP_000447;
CC Name=4;
CC IsoId=Q9H6X2-4; Sequence=VSP_000448; VSP_000449;
CC -1- TISSUE SPECIFICITY: Highly expressed in tumor endothelial cells
CC but not in normal endothelial cells.
CC -1- DOMAIN: Binding to PA seems to be effected through the VMA domain.
CC -1- SIMILARITY: Belongs to the ATR family.
CC -1- SIMILARITY: Contains 1 VMA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AF279145; AAK52094.1; -;
CC EMBL; AF421380; AAL26496.1; -;
CC EMBL; AK025429; BAB15128.1; ALT_INIT.
CC EMBL; AK011463; BAA91707.1; ALT_FRAME.
CC EMBL; BC012074; AAI12074.1; -;
CC GeneW; HGNC:21014; ANTXRL.
CC H-InvDB; HIX0002125; -;
CC MIM; 606410; -;
CC InterPro; IPR008400; Anth_Ig.
CC InterPro; IPR008399; Ant_C.
CC InterPro; IPR002035; VWF_A.
CC Pfam; PF05587; Anth_Ig; 1.
CC Pfam; PF05586; Ant_C; 1.
CC Pfam; PR00982; VMA; 1.
CC SMART; SM00327; VMA; 1.
CC DR PROSITE; PS50234; VWF; 1.
CC KW Alternative splicing; Glycoprotein; Receptor; Signal; Transmembrane.
CC FT CHAIN 1 32
CC FT SIGNL 1 32
CC FT DOMAIN 33 321
CC FT TRANSMEM 322 342
CC FT DOMAIN 343 564
CC FT DOMAIN 44 215
CC FT DOMAIN 360 368
CC FT CARBOHYD 166 166
CC FT CARBOHYD 184 184
CC FT CARBOHYD 262 262
CC FT VARSPPLIC 365 368
CC FT VARSPPLIC 369 564
CC FT VARSPPLIC 268 297
CC FT VARSPPLIC 298 564
CC FT VARSPPLIC 319 333
CC FT VARSPPLIC 334 564
CC SQ SEQUENCE 564 AA; 62789 MW; B118A00D5DF2233 CRG64;
CC Query Match 100.0%; Score 943; DB 1; Length 564;
CC Best Local Similarity 100.0%; Pred. No. 9, 2e-73;
CC Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 1 GPDYFILLKSGSVLHNNMEIYFVEQLAHKFI SPOLRMSFIYSTRGTTMLKLTDBREQ 60
CC DB 42 GPDYFILLKSGSVLHNNMEIYFVEQLAHKFI SPOLRMSFIYSTRGTTMLKLTDBREQ 101
CC QY 61 IROGLEELQKVLPGSDTYNHEGFERASEQIYYENRGYRTASYIALTDGELHEDLFFYS 120
CC DB 102 IROGLEELQKVLPGSDTYNHEGFERASEQIYYENRGYRTASYIALTDGELHEDLFFYS 161
CC QY 121 BREANRSRLGALIVCVGVKDFNETOLARIADSKDHVPVNDGFQALOGIHSILKKSCT 180
CC DB 162 BREANRSRLGALIVCVGVKDFNETOLARIADSKDHVPVNDGFQALOGIHSILKKSCT 221
CC QY 181 E 181
CC DB 222 E 222
CC
CC RESULT 2
CC ATRL_MOUSE STANDARD; PRT; 562 AA.

AC Q9CZ52.2; (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Anthrax toxin receptor 1 precursor (tumor endothelial marker 8).
 GN Name=Antxr1; Synonyms=Atlr, Tem8;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=21443268; PubMed=11559528;
 RA Carson-Walter E.B., Watkins D.N., Nanda A., Vogelstein B.,
 RA Kinzler K.W., St Croix B.;
 RT "Cell surface tumor endothelial markers are conserved in mice and
 humans.";
 RL Cancer Res. 61:6649-6655 (2001).
 RN [2]
 RP SEQUENCE OF 68-562 FROM N.A. (ISOFORM 2).
 RC STRAIN=C57BL/6J; TISSUE=Embryo.
 RX MEDLINE=22354683; PubMed=12466651; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nakafuku I., Oono N., Saito R., Suzuki H., Yamana K.I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Haegawa Y., Nogami A., Schonbach C., Gotohori T.,
 RA Balderelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schmitt L.M., Kanapin A., Matcoda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Busic V., Chochia S., Corbani L.E., Cousins S.,
 RA Dalla B., Dargatz T.A., Fletcher C.F., Forrest A., Frezer K.S.,
 RA Gaesteland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grifmond S., Guerlinch S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedziarski R.M., King B.L.,
 RA Konoaga A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Matsuoka L., Marchionni L., McKenzie L., Miki H.,
 RA Nagasubramanian T., Numa K., Okita J., Pavan W.J., Pereira G., Pesole G.,
 RA Petrovsky N., Piliat R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Savelle A., Schneider C., Sempke C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Vercarlo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilmink L.G., Wyshewski-Borja A., Yamaoka S., Yang I., Yang L.,
 RA Yuan Z., Zavalon M., Zhu Y., Zimmer A., Carlini P., Hayatsu N.,
 RA Hirozane-Kobayashi K., Kono H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Aizawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imocani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sakai D., Shibata K., Shingawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573 (2002).
 CC -1 FUNCTION: Cellular role is not yet known.
 CC -1 SUBUNIT: Binds to the protective antigen (PA) of Bacillus
 CC anthracis (by similarity).
 CC -1 SUBCELLULAR LOCATION: Type I membrane protein (Probable).
 CC -1 ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q9CZ52-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q9CZ52-2; Sequence=VSP_000450;
 CC Note=No experimental confirmation available;
 CC -1 DOMAIN: Binding to PA seems to be effected through the VWA domain
 CC (by similarity).
 CC -1 SIMILARITY: Belongs to the ATR family.
 CC -1 SIMILARITY: Contains 1 VWA domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)

Query	Match	Best Local Similarity	98.9%	Score 934	DB 1	Length 562
Db	1	98.9% <td>Score 934 <td>DB 1 <td>Length 562 <td></td> </td></td></td>	Score 934 <td>DB 1 <td>Length 562 <td></td> </td></td>	DB 1 <td>Length 562 <td></td> </td>	Length 562 <td></td>	
Qy	61	98.9% <td>Score 934 <td>DB 1 <td>Length 562 <td></td> </td></td></td>	Score 934 <td>DB 1 <td>Length 562 <td></td> </td></td>	DB 1 <td>Length 562 <td></td> </td>	Length 562 <td></td>	
Db	40	98.9% <td>Score 934 <td>DB 1 <td>Length 562 <td></td> </td></td></td>	Score 934 <td>DB 1 <td>Length 562 <td></td> </td></td>	DB 1 <td>Length 562 <td></td> </td>	Length 562 <td></td>	
Qy	121	98.9% <td>Score 934 <td>DB 1 <td>Length 562 <td></td> </td></td></td>	Score 934 <td>DB 1 <td>Length 562 <td></td> </td></td>	DB 1 <td>Length 562 <td></td> </td>	Length 562 <td></td>	
Db	160	98.9% <td>Score 934 <td>DB 1 <td>Length 562 <td></td> </td></td></td>	Score 934 <td>DB 1 <td>Length 562 <td></td> </td></td>	DB 1 <td>Length 562 <td></td> </td>	Length 562 <td></td>	
Qy	181	98.9% <td>Score 934 <td>DB 1 <td>Length 562 <td></td> </td></td></td>	Score 934 <td>DB 1 <td>Length 562 <td></td> </td></td>	DB 1 <td>Length 562 <td></td> </td>	Length 562 <td></td>	
Db	220	98.9% <td>Score 934 <td>DB 1 <td>Length 562 <td></td> </td></td></td>	Score 934 <td>DB 1 <td>Length 562 <td></td> </td></td>	DB 1 <td>Length 562 <td></td> </td>	Length 562 <td></td>	

RT membrane matrix assembly, cell cycle progression, cellular
RT differentiation and G-protein signaling." ,
RL J. Cell Sci. 114:2755-2773 (2001) .
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND INTERACTION WITH ANTHRAX TOXIN.
RC TISSUE=Placenta;
RX MEDLINE=22606610; Pubmed=12700348; DOI=10.1073/pnas.0431098100;
RA Scobie H.M., Rainey G.J.A., Bradley K.A., Young J.A.T.;
RT "human capillary morphogenesis protein 2 functions as an anthrax toxin
receptor." ;
RL Proc. Natl. Acad. Sci. U.S.A. 100:5170-5174 (2003) .
RN [3]
RP SEQUENCE OF 78-489 FROM N.A. (ISOFORM 3), AND SEQUENCE FROM N.A.
RN (ISOFORM 4) .
RC TISSUE=Synovial cell;
RX Pubmed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Makatsutsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohbayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Nagahara K., Murakami K., Iwata T., Iwayanagi T., Wagatsuma M.,
RA Shiratori A., Sudo H., Hosoi T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RA Yamazaki M., Nimomiya K., Iehiboshi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hizoka S., Chiba Y.,
RA Iehida S., Ono Y., Takiguchi S., Watanabe S., Yoshida M., Houcha T.,
RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Togiyama S., Komai F., Hara R., Takeuchi K., Arita M.,
RA Imose N., Masashiro K., Yuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki M., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Taniguchi A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujitawa T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Omori Y.,
RA Kababata A., Hikiji T., Kobayashi N., Inagaki H., Ikema Y., Okamoto S.,
RA Oktani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsunura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuno Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isegaki T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
cDNAs." ;
RL Nat. Genet. 36:40-45 (2004) .
RN [4]
RP FUNCTION: Cellular role is not yet known.
CC -1- SUBUNIT: Binds to the protective antigen (PA) of Bacillus
CC anthracis in a divalent cation-dependent manner, with the
CC following preference: calcium > manganese > magnesium > zinc.
CC Seems to bind to collagen type IV and laminin.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (isoforms 1 and 2).
CC secreted (isoform 3). Isoform 1 is expressed at the cell surface
CC while isoform 2 is predominantly expressed within the endoplasmic
CC reticulum and not at the plasma membrane.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Name=1;
CC NameId=P58335-1, Sequence=Displayed;
CC Name=2;
CC NameId=P58335-2, Sequence=VSP_008343;
CC Name=3;
CC NameId=P58335-3, Sequence=VSP_008344, VSP_008345;
CC Note=No experimental confirmation available;
CC Name=4;
CC NameId=P58335-4, Sequence=VSP_008346;
CC Note=No experimental confirmation available;
CC -1- TISSUE SPECIFICITY: Expressed in colon, heart, kidney, lung,
CC liver, peripheral blood leukocytes, placenta, skeletal muscle,
CC small intestine and spleen.
CC -1- DOMAIN: Binding to PA seems to be effected through the VWA domain.
CC -1- SIMILARITY: Belongs to the ATR family.

CC	-!- SIMILARITY: Contains 1 VMPA domain.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.ebi.ac.uk/announce/
CC	or send an email to license@ebi.ac.uk).
CC	-----
DR	EMBL; AY040326; AAK77222.1; -
DR	EMBL; AY233452; AAP04016.1; -
DR	EMBL; AK055636; BAB70976.1; ALT_INIT.
DR	EMBL; AK091721; BAC03731.1; -
DR	GeneW; HGNC:21732; ANTXR2.
DR	MIM; 608041; -
DR	InterPro; IPR008399; Ant C.
DR	InterPro; IPR020355; VME_A.
DR	Pfam; PF05586; Ant_C_1.
DR	Pfam; PF00092; VMA_1.
KW	Alternative splicing; Glycoprotein; Receptor; Signal; Transmembrane.
FT	SIGNAL 1 33
FT	CHAIN 34 489
FT	DOMAIN 34 318
FT	TRANSMEM 319 341
FT	DOMAIN 342 489
FT	DOMAIN 44 213
FT	CARBOHYD 250 250
FT	CARBOHYD 260 260
FT	VARSPLIC 213 315
FT	/FTID=VSP_008343.
FT	/FTID=VSP_008344.
FT	VLQGVSVFNGKSVISGLIVTATCGNGINAI -> WGLT
FT	VTQAGVWMHDLTHCTFGLSGSDPPTSAS (in isoform
FT	3).
FT	/FTID=VSP_008344.
FT	Missing (in isoform 3).
FT	/FTID=VSP_008345.
FT	VCIMECEKELEYA -> GRCLNFSRVPSQ (in isoform
FT	4).
FT	/FTid=VSP_008346.
FT	P -> A (in Ref. 3; BAC03731).
FT	B9F679DB75BE2B7 CRC64;
SO	SEQUENCE 489 AA; 53692 MW; B9F679DB75BE2B7 CRC64;
Query Match	57.5%; Score 542; DB 1; Length 489;
Best Local Similarity	59.4%; Pred. No. 2,7e-38;
Matches 107,	Conservative 35; Mismatches 36; Indels 2; Gaps 1;
QY	2 FDLFFIIDKSGSVLAHNNETIYPVEQLAHKFISQLMNSFLVSTRGTTLMKLTEDEBOI 61
Db	43 FDLFFVLDKSGSVANNMIETIYNFQQALAFVSPSEMLSFVFSSOATITLIPLTGDDGKI 102
QY	62 ROGEEELOKVLPGGDTVMHGGEFAASAHOIYYENRQGRTSVITALTDGELHEDLFPSYE 121
Db	103 SKGLEEDLKRFSPVEGTIIHERLKLANBOI--QRAGKLTSTIIALTGDKLDGLVPSTAE 160
QY	122 REARSRKDLGIATYCVGVKDFNETQTOLRIADSKOHVPNVNDGFQALOGIISHSIKKSCIE 181
Db	161 KEATISHSLSGSAYCIGVDLFEQAQLERINDSKEQVFPVGGFPAALGIIINSILAQCTE 220
RESULT 4	
O6DPFX2	PRELIMINARY; PRT; 487 AA.
AC	O6DPFX2;
DT	25-OCT-2004 (TREMBLrel. 28, Created)
DT	25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT	25-OCT-2004 (TREMBLrel. 28, last annotation update)
DE	Anthrax toxin receptor 2.
GN	Name=Antxr2;
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurgnathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Eye;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strassberg R.L., Feinold E.A., Grouse L.H., Derge J.G.,
RA Krausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuller G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Datchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Urdin T.B., Tohiyuki S., Carninci P., Prange C.T.,
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.U., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richerds S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Phelan J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bonfield G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzyzaniak M.I., Skalek U., Small D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Eye;
RA Strassberg R.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC076595; AAI76595.1; -
DR GO: GO:0016021; C:Integral to membrane; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR InterPro: IPR008400; AnCh_Ig.
DR InterPro: IPR008399; Ant_C.
DR InterPro: IPR002035; VWF_A.
DR Pfam: PF05587; AnCh_Ig; I.
DR Pfam: PF05586; Ant_C; 1.
DR Pfam: PF00092; VWA; 1.
DR SMART: SM00327; VWA; 1.
DR PROSITE: PSS0234; VWFA; 1.
KM Receptor.
SQ SEQUENCE 487 AA; 53184 MW; 61A00D60BC8DE69 CRC64;
Query Match 57.3%; Score 540; DB 2; Length 487;
Best local similarity 59.4%; Pred. No. 4e-38;
Matches 107; Conservative 32; Mismatches 39; Indels 2; Gaps 1;
QY 2 FDYFLIDSGSVLHNNELIYFVEQLAHKFISPOLMSFIVSTRGTTMLKLTEDREOI 61
DB 43 FDLYFLIDSGSVLHNNELIYFVEQLAHKFISPOLMSFIVSTRGTTMLKLTEDREOI 102
QY 62 RQLEBELQKLVPCGDTYTMAGGFERASQIYENKQYRTASVITLTDGELHEDLPFYSE 121
DB 103 GKLEBELQKLVPCGDTYTMAGGFERASQIYENKQYRTASVITLTDGELHEDLPFYSE 160
QY 122 REANRSRDGLAIYVCGVDENFETOLARIADSKDHPVAVDQFQALGIIHSLIKSCIE 181
DB 161 NEAKRSRSLGASVYCGVLDPEQQLERLADSKDQVFPVKGFGQALGIIHSLIKSCIE 220
RESULT 5
Q88VM2 PRELIMINARY; PRT; 641 AA.
AC Q88VM2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Mus musculus adult male testis cDNA, RIKEN full-length enriched
DE library, clone:4933430J11 product:hypothetical Prolin-rich region/von
DE Willebrand factor type A domain containing protein, full insert
DE sequence.
OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Testis;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning";
RL Meth. Enzymol. 303:19-44(1999).
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection";
RL Nature 409:685-690(2001).
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Testis;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573(2002).
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Testis;
RX MEDLINE=20493374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes";
RL Genome Res. 10:1617-1630(2000).
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Testis;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nageoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Niehi K., Kiteunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer";
RL Genome Res. 10:1757-1771(2000).
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Testis;
RA Aachisi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hamagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Horii F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kishihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ono M., Ohnato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakata N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK077206; BAC36683.1; -
DR GO: GO:0016021; C:Integral to membrane; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR InterPro: IPR008400; AnCh_Ig.
DR InterPro: IPR002035; VWF_A.
DR Pfam: PF05587; AnCh_Ig; I.
DR Pfam: PF00092; VWA; 1.
DR PRINTS: PR00453; VWFADOMAIN.
DR SMART: SM00327; VWA; 1.
DR PROSITE: PSS0234; VWFA; 1.

KW Hypothetical protein.
SQ SEQUENCE 641 AA; 70415 MW; 199E300730BC85E3 CRC64;
Query Match 39.8%; Score 375; DB 2; Length 641;
Best Local Similarity 43.9%; Pred. No. 9e-24;
Matches 79; Conservative 31; Mismatches 70; Indels 0; Gaps 0;
QY 2 FDLVFLDKSGSVLHNNELIYFVEQLAKKFIPOKMSFIVSTRGTMLKLTEDREQ 61
DB FDLVFLDKSGSVADNMIHLYSFAGLVKKFTNPENLISITVSTAEVLLPLTDSKEI 134
QY 62 RQGLEELQKVLPGDVTWHEGFERASQIYYENNQGVRTSAVIALTDGELHEDLPFYSE 121
DB 135 NKGILVTKSVIPQGLTMQKGRKANQIKRSTIGRIIVNSVIALTDGILLKPYIDTM 194
QY 122 REANRSDLAIVYCVGKDFNETOLARIADSKDHVPVNDFOALOGIIHSILKSCIE 181
DB 195 EBAKKARMGALVYTVGVFMYSKQVLNIGDPDRCGVDEGFALGEGVDPLTSKSCIE 254
RESULT 6
Q9BP08 PRELIMINARY; PRT; 1332 AA.
AC Q9BP08;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE Integrin alpha H1 precursor.
GN Name=H1RGAL;
OS Halocynthia roretzi (Sea squirt).
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Scolidobranchia; Pyuridae; Halocynthia.
OX NCBI_TaxId=7729;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hemocyte;
RA MEDLINE=21103187; Pubmed=11160215;
RX Miyazawa S., Azumi K., Nonaka M.,
RT "Cloning and characterization of integrin alpha subunits from the
RT solitary ascidian, Halocynthia roretzi.",
RL J. Immunol. 166:1710-1715 (2001).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
DR EMBL; AB048261; BAB21479.1; -.
DR HSSP; P11215; 1BHQ.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0008305; C:Integrin complex; IEA.
DR GO; GO:0005515; F:Protein binding; IEA.
DR GO; GO:0007160; P:Cell-matrix adhesion; IEA.
DR GO; GO:0007229; P:Integrin-mediated signaling pathway; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF00092; VMA; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VMA; 1.
DR PROSITE; PS50234; VWFA; 1.
KW Cell adhesion; Integrin; Signal; Transmembrane.
FT SIGNAL 1 30 Potential.
FT CHAIN 31 1332 Integrin alpha H1.
SQ SEQUENCE 1332 AA; 145851 MW; 0D9108D2B05CFB4 CRC64;
Query Match 15.5%; Score 146; DB 2; Length 1332;
Best Local Similarity 24.5%; Pred. No. 0.0011;
Matches 52; Conservative 34; Mismatches 60; Indels 66; Gaps 7;
QY 1 GFLVFLDKSGSVLHNNELIYFVEQLAKKFIPOKMSFIVSTRGTMLKLTEDREQ 60
DB GAVVFLFLDKSGSVGKDFDKVKKVKNIT-----AKLDIGKEI 241
QY 61 IROGLEELQKVLPGDVTWHEGFERASQIYYENNQGVRT----- 100

DB 242 VRVGVQYSHYVSGKSIINKQYITTEISIGFELDNFENAVDRI---QLQGYTTYGRA 298
QY 101 -----ASVIALTDGELHEDLPFYSSREANRSSDLAIVYCVGKDFNET 145
DB 299 LQKVRIDFDPAVIGNKQVLLLTLDGQAKDKILP--NANRLNKGATPAVGGEYDIS 356
QY 146 QLARIA---DSKDHVPVNDFOALOGIIHSI 174
DB 357 ELKLIASGTDSTDRVFTVD-FGELDSIVKSL 387
RESULT 7
ITAD_RAT
ID ITAD_RAT STANDARD; PRT; 1161 AA.
AC Q9QYE7;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Integrin alpha-D precursor.
GN Name=Itgad;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA O'Brien M.M., VanderVlieten M., Kilgannon P.D., Dietech G.,
RA Gallatin W.M.;
RT "Cloning of rat alpha D, a novel beta 2 integrin."
RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Integrin alpha-D/beta-2 is a receptor for ICM3 and VCAM1. May play a role in the atherosclerotic process such as clearing lipoproteins from plaques and in phagocytosis of blood-borne pathogens, particulate matter, and senescent erythrocytes from the blood (By similarity).
CC -1- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-D associates with beta-2 (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -1- DOMAIN: The integrin I-domain (Inser1) is a VWFA domain. Integrins with I-domains do not undergo protease cleavage.
CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
CC -1- SIMILARITY: Contains 1 VWFA domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
DR EMBL; AF021334; AAP21241.1; -.
DR HSSP; P11215; 1BHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_alpha; 1.
DR Pfam; PF00092; VMA; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VMA; 1.
DR PROSITE; PS50242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWFA; 1.
KW Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
FT Repeat; Signal; Transmembrane.
FT SIGNAL 1 19 Potential.
FT CHAIN 20 1161 Integrin alpha-D.
FT DOMAIN 20 1100 Extracellular (Potential).
FT TRANSMEM 1101 1121 Potential.
FT DOMAIN 1122 1161 Cytoplasmic (Potential).


```

FT REPEAT 34 87 FG-GAP 1.
FT REPEAT 88 7 FG-GAP 2.
FT DOMAIN 152 334 VFMA.
FT REPEAT 352 402 FG-GAP 3.
FT REPEAT 403 454 FG-GAP 4.
FT REPEAT 456 517 FG-GAP 5.
FT REPEAT 519 577 FG-GAP 6.
FT REPEAT 582 634 FG-GAP 7.
FT CA_BIND 467 475 Potential.
FT CA_BIND 531 539 Potential.
FT CA_BIND 594 602 Potential.
FT SITE 1126 1130 GFPKR motif.
FT DISULFID 69 76 By similarity.
FT DISULFID 108 126 By similarity.
FT DISULFID 656 711 By similarity.
FT DISULFID 769 775 By similarity.
FT DISULFID 845 860 By similarity.
FT DISULFID 993 1017 By similarity.
FT DISULFID 1022 1027 By similarity.
FT CARBOHYD 61 61 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 89 89 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 245 245 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 393 393 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 696 696 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 734 734 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 784 784 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 907 907 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 936 936 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1045 1045 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 1161 AA; 126600 MW; 2258491A984A705E CRC64;

```

```

Query Match 15.0%; Score 141.5; DB 1; Length 1161;
Best Local Similarity 27.7%; Pred. No. 0.0023;
Matches 53; Conservative 38; Mismatches 75; Indels 25; Gaps 10;

```

```

QY 3 DLYFIIDKSSGV-LHHNNEITYFPEQLAHKFIISPOLMSFVSTRTGTTMLKLT----- 56
DB 152 DIAFLIDGSGISINORDPAQKMDPKALMGFASSTLFSLMQYSNLIKHTFTFEFNIL 211
QY 57 DREQIROGLELOKVLPG--GDTYMHGEPERASQIYYENRQGYRTA-SVITALTDELHED 115
DB 212 DPGSLVDPIYQLQ-----GLITRTATGIRITWEEIFHSKNSRKSAAKILVIITDQKTRD 266
QY 116 LFPEYSR--REANRSRDGAIVYCVGV--D-FNE--TOLARI-----ADSKDHVFPVNDGFOA 166
DB 267 PLEYSDVTPADAXA---GIRYALGVGDARFOEPALMELMTIGSAPRQDHVFKYGN-FAA 322
QY 167 LQGIHISILKK 177
DB 323 LRISIORLOERK 333

```

```

RESULT 8
Q8T6U5 PRELIMINARY; PRT; 441 AA.
ID Q8T6U5;
AC Q8T6U5;
DT 01-JUN-2002 (TREMblrel. 21, Last Created)
DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)
DE Proximal thread matrix protein 1 variant a.
OS Mytilus edulis (Blue mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_TaxId=6550;
RN [1]
RS SEQUENCE FROM N.A.
RA MEDLINE=22313036; PubMed=12425661; DOI=10.1021/bm0255903;
RX Sun C., Lucas J.M., Waite J.H.;
RT "Collagen-binding matrix proteins from elastomeric exteroorganismic
RT byssal fibers.";
RL Biomacromolecules 3:1240-1248 (2002).
DR EMBL; AF414454; AAL83537.1; -
DR GO; GO:0005198; F:structural molecule activity; IEA.

```

```

DR InterPro: IPR002035; VMA_A.
DR Pfam; PF00092; VMA; 2.
DR PRINTS; PR00453; VMAFAMAIN.
DR SMART; SM00327; VMA; 2.
DR PROSITE; PS50234; VMAF; 2.
KM Matrix protein.
SQ SEQUENCE 441 AA; 47543 MW; 881DBBD36B891D2B CRC64;

```

```

Query Match 14.7%; Score 139; DB 2; Length 441;
Best Local Similarity 25.1%; Pred. No. 0.0012;
Matches 49; Conservative 46; Mismatches 74; Indels 26; Gaps 12;

```

```

QY 3 DLYFIIDKSSGV-LHHNNEITYFPEQLAHKFIISPOLMSFVSTRTGTTMLKLT 55
DB 241 DIAVFDPASSINANNNNYGLMKDIFVDRFNKTPGPGTQPAVVTFAADRATKQFGLK 300
QY 56 E--DREQIROGLELOKVLPG--GDTYMHGEPERASQIYYENRQ--VRTASVITALT 108
DB 301 DYSSKAIRKGAID--KVTSIIIGQTALIGDLENARLEV-FPRNNGGGRREVQKVILLT 356
QY 109 DGBL--HEDLFYSEREANRSRDGAIVYCVGV--KDFNETOLARIADSKDHVFPVNDGFO 165
DB 357 DGQNGHKS----PEHSSILRKRGVIVTALGVGTGFLKSELINIASSEYVF-TTSSFN 411
QY 166 ALQGIHISILKSGCT 180
DB 412 KLSKIMENVVYLACM 426

```

RESULT 9

```

Q8T5C3 PRELIMINARY; PRT; 444 AA.
ID Q8T5C3;
AC Q8T5C3;

```

```

DT 01-JUN-2002 (TREMblrel. 21, Last Created)
DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)
DE Proximal thread matrix protein 1b.
OS Mytilus edulis (Blue mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_TaxId=6550;
RN [1]
RS SEQUENCE FROM N.A.
RA Sun C., Lucas J.M., Waite J.H.;
RT "Collagen-binding matrix proteins from elastomeric exteroorganismic
RT byssal fibers.";
RL Biomacromolecules 3:1240-1248 (2002).
DR EMBL; AY053390; AAL17973.1; -
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR002035; VMAF_A.
DR Pfam; PF00092; VMA; 2.
DR PRINTS; PR00453; VMAFAMAIN.
DR SMART; SM00327; VMA; 2.
DR PROSITE; PS50234; VMAF; 2.
KM Matrix protein.
SQ SEQUENCE 444 AA; 47615 MW; D2C605347450C931 CRC64;

```

```

Query Match 14.7%; Score 139; DB 2; Length 444;
Best Local Similarity 25.1%; Pred. No. 0.0012;
Matches 49; Conservative 46; Mismatches 74; Indels 26; Gaps 12;

```

```

QY 3 DLYFIIDKSSGV-LHHNNEITYFPEQLAHKFIISPOLMSFVSTRTGTTMLKLT 55
DB 244 DIAVFDPASSINANNNNYGLMKDIFVDRFNKTPGPGTQPAVVTFAADRATKQFGLK 303
QY 56 E--DREQIROGLELOKVLPG--GDTYMHGEPERASQIYYENRQ--VRTASVITALT 108
DB 304 DYSSKAIRKGAID--KVTSIIIGQTALIGDLENARLEV-FPRNNGGGRREVQKVILLT 359
QY 109 DGBL--HEDLFYSEREANRSRDGAIVYCVGV--KDFNETOLARIADSKDHVFPVNDGFO 165
DB 360 DGQNGHKS----PEHSSILRKRGVIVTALGVGTGFLKSELINIASSEYVF-TTSSFN 414

```

```
Qy      166 ALOGIHSILKSGCI 180
      | : : : : | :
Db      415 KLSKIMENVVKLACM 429

RESULT 10
Q875C2
ID Q875C2 PRELIMINARY; PRT; 453 AA.
AC Q875C2;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Proximal thread matrix protein 1
OS Mytilus galloprovincialis (Mediterranean mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_TaxID=29158;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22313036; PubMed=12425661; DOI=10.1021/bm0255903;
RA Sun C., Lucas J.M., Waite J.H.;
RT "Collagen-binding matrix proteins from elastomeric extraorganic
RT byssal fibers.";
RL Biomacromolecules 3:1240-1248(2002).
DR HSSP; P20701; IMCN.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR02035; VWF_A.
DR Pfam; PF00092; VMA; 2.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00327; VMA; 2.
DR PROSITE; PS50234; VWF_A; 2.
DR Matrix protein.
KW SEQUENCE 453 AA; 48784 MW; D60497F5C0C51EED CRC64;
SQ

Query Match 14.4%; Score 136; DB 2; Length 453;
Best Local Similarity 25.1%; Pred. No. 0.0023;
Matches 49; Conservative 44; Mismatches 76; Indels 26; Gaps 12;

Qy      3 DLVFIIDKSGSV---LHMNEIYFVEQLAHKF--ISPO-LRMSFIVFSTRTGLTKLT 55
      | : : : : | : : : : | : : : : | : : : : | : : : : |
Db     253 DIAFVFPASSINANNNNVQLMKNFMKDIYDRNKTGPOTGPAVVTADRAIKRGK 312
Qy      56 E--DREQIRGLLEQLVLP--GDTYMHGFEFASQIYENRQ--YRTASVITALT 108
      | : : : : | : : : : | : : : : | : : : : | : : : : |
Db     313 DYSSKADIKGAID---KVSPIIGQTALGDLENARLEV-FPNNGGGRREVQKVILIT 368
Qy      109 DGEI--HEDLFPYSEERANRSDIGALVYCVG-KDENEQTARLADSKDHVPYNDG 165
      | : : : : | : : : : | : : : : | : : : : | : : : : |
Db     369 DQNNNGHKS---PEHSSSLRKGVVIVAGVGTGFLKSELINIASSEEVF-TTISPD 423
Qy      166 ALOGIHSILKSGCI 180
      | : : : : | :
Db     424 KLSKIMENVVKLACM 438

RESULT 11
Q04588
ID Q04588 PRELIMINARY; PRT; 724 AA.
AC Q04588;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Major antigen homologous sequence (emp100).
OS Eimeria maxima.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae;
OC Eimeria.
OX NCBI_TaxID=5804;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=3149203; PubMed=8426611; DOI=10.1016/0166-6851(93)90255-V;
RA Paganontes L.E., Hug D., Huembelin M., Weber G.;
```

```
RT      "Sequence of a major Eimeria maxima antigen homologous to the Eimeria
RT tenella microneme protein Exp100.";
RL Mol. Biochem. Parasitol. 57:171-174(1993).
DR EMBL; M99058; AAA29076.1; -.
DR PIR; A48569; A48569.
DR HSSP; P07996; 1LSL.
DR GO; GO:0008083; F:growth factor activity; IEA.
DR InterPro; IPR000762; PTN_MK.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF00090; TSP_1; 5.
DR Pfam; PF00092; VMA; 1.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00209; TSP1; 6.
DR SMART; SM00327; VMA; 1.
DR PROSITE; PS50092; TSP1; 5.
DR PROSITE; PS50234; VWF_A; 1.
DR SEQUENCE 724 AA; 75808 MW; AC2A0E7A346A7E9E CRC64;
SQ

Query Match 14.4%; Score 136; DB 2; Length 724;
Best Local Similarity 27.0%; Pred. No. 0.004;
Matches 53; Conservative 36; Mismatches 79; Indels 28; Gaps 12;

Qy      3 DLVFIIDKSGSV-LHMNEIYFVEQLAHKF-ISP-OLRMSFIVFSTRTGLTKLTEDRE 59
      | : : : : | : : : : | : : : : | : : : : | : : : : |
Db     47 DWLVVDDESGISITSYGKRSISNPFAGTMPLSPDVGVLVTFTGSATVRDLSRSRA 106
Qy      60 QIRQGLEELQKLP--GDTYMHGFEFASQIYENRQYR--TASVITATDGEIHED 115
      | : : : : | : : : : | : : : : | : : : : | : : : : |
Db     107 QNADLLAAAKKLPVAASTYTHLGLAKA-EELIFSGKGRDNAPMILVMTDGA----- 161
Qy      116 LFFYSERE-----AKSRDLGALVYCVG-KDNEQTARL--DSKHV-PP--VNDG 163
      | : : : : | : : : : | : : : : | : : : : | : : : : |
Db     162 ---SSRRSQTSAAKLRNRGVIVAGVGVNSAECSRIAGCDTVECPRYQSN 217
Qy      164 FQALOGIHSILKSGC 179
      | : : : : | :
Db     218 WGVSSQINGITIKAC 233

RESULT 12
CALC_MOUSE
ID CALC_MOUSE STANDARD; PRT; 3119 AA.
AC Q60847; P70322;
DT 15-JUL-1998 (Rel. 36, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Collagen alpha 1(XII) chain precursor.
GN Name=Coll2a1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP PARTIAL SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING (ISOFORMS XIIA-1 AND
RP XIIB-1).
RP STRAIN=C57BL/6J; TISSUE=Skin fibroblast;
RX MEDLINE=96170761; PubMed=8601036;
RA Boehme K., Li Y., Oh P.S., Olsen B.R.;
RT "Primary structure of the long and short splice variants of mouse
RT collagen XII and their tissue-specific expression during embryonic
RT development.";
RL Dev. Dyn. 204:432-445(1995).
RN [2]
RP PARTIAL SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING (ISOFORMS XIIA-2
RP AND XIIB-2).
RP STRAIN=C57BL/6J; TISSUE=Skin fibroblast;
RX MEDLINE=99348349; PubMed=10419532; DOI=10.1074/jbc.274.31.22053;
RA Kania A.M., Reichenberger E., Baur S.T., Karimbux N.Y., Taylor R.W.,
RA Olsen B.R., Nishimura I.;
RT "Structural variation of type XII collagen at its carboxyl-terminal
RT NCI domain generated by tissue-specific alternative splicing.";
RL J. Biol. Chem. 274:22053-22059(1999).
```

CC -/- FUNCTION: Type XII collagen interacts with type I collagen-
 CC containing fibrils, the COL1 domain could be associated with the
 CC surface of the fibrils, and the COL2 and NC3 domains may be
 CC localized in the pericellular matrix (By similarity).
 CC -/- SUBUNIT: Trimer of identical chains each containing 190 kDa of
 CC nontriple-helical sequences (By similarity).
 CC -/- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=4;
 CC Comment=The final tissue form of collagen XII may contain
 CC homotrimers or any combination of the various isoforms;
 CC Name=XIIA-1;
 CC IsoId=Q060847-1; Sequence=Displayed;
 CC Name=XIIA-2; Synonyms=ERK1;
 CC IsoId=Q060847-2; Sequence=VSP_001151, VSP_001152;
 CC Name=XIIB-1;
 CC IsoId=Q060847-3; Sequence=VSP_001150;
 CC Name=XIIB-2;
 CC IsoId=Q060847-4; Sequence=VSP_001150, VSP_001151, VSP_001152;
 CC -/- TISSUE SPECIFICITY: Highest expression in tendon, perichondrium,
 CC skin, cornea, sclera, blood vessels, and pericardium.
 CC -/- DEVELOPMENTAL STAGE: The long NC3 XIA isoforms are predominant at
 CC early stages (ED7 and 11); at later stages of development (ED15
 CC and 17) the short NC3 XIIB forms become the major forms. As the
 CC short NC3 forms become the major product, the long splice variant
 CC continues to be expressed in several tissues, even after birth.
 CC The long NC1 isoform, XIA-1 and XIIB-1, peak in 15-day old
 CC embryos and decrease in 17-day old ones. The expression of the
 CC short NC1 form XIIB-2 remains constant throughout late stages of
 CC embryonic development (ED15 and ED17).
 CC -/- PTM: The triple-helical tail is stabilized by disulfide bonds at
 CC each end (By similarity).
 CC -/- PTM: Prolines at the third position of the tripeptide repeating
 CC unit (G-X-Y) are hydroxylated in some or all of the chains (By
 CC similarity).
 CC -/- PTM: O-glycosylation of isoform XIIA-2, glycosaminoglycan of
 CC chondroitin-sulfate type (By similarity).
 CC -/- SIMILARITY: Belongs to the fibril-associated collagens with
 CC interrupted helices (FACIT) family.
 CC -/- SIMILARITY: Contains 18 fibronectin type III domains.
 CC -/- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
 CC -/- SIMILARITY: Contains 4 VMPA domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by, and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; U25652; AAA99719.1; ALT_SEQ.
 CC EMBL; U57095; AAB07047.1; -.
 CC HSSP; P18614; LMHP.
 CC MGD; MG188448; COL12a1.
 CC InterPro; IPR008160; Collagen.
 CC InterPro; IPR008985; Cona1like_1ec_g1.
 CC InterPro; IPR003961; FN_III.
 CC InterPro; IPR008957; FN_III-like.
 CC InterPro; IPR003129; TSP_N.
 CC InterPro; IPR002035; VMP_A.
 CC Pfam; PF01391; Collagen_5.
 CC Pfam; PF00041; fn3_18.
 CC Pfam; PF02210; TSP_N_1.
 CC Pfam; PF00092; VMA_4.
 CC PRINTS; PR00453; VMPADOMAIN.
 CC SMART; SM00060; FN3_18.
 CC SMART; SM00210; TSPN_1.
 CC SMART; SM00327; VMA_4.
 CC PROSITE; PS50853; FN3_18.
 CC PROSITE; PS50234; VMPA_4.
 CC Alternative splicing; Cell adhesion; Collagen; Extracellular matrix;
 CC Glycoprotein; Hydroxylation; Repeat; Signal; Structural protein.
 CC SIGNAL 1 24 Potential.

FT	CHAIN	25	3119	Collagen alpha 1(XII) chain.
FT	DOMAIN	25	112	Fibronectin type-III 1.
FT	DOMAIN	140	316	VMPA 1.
FT	DOMAIN	333	422	Fibronectin type-III 2.
FT	DOMAIN	444	620	VMPA 2.
FT	DOMAIN	635	723	Fibronectin type-III 3.
FT	DOMAIN	726	814	Fibronectin type-III 4.
FT	DOMAIN	817	905	Fibronectin type-III 5.
FT	DOMAIN	908	997	Fibronectin type-III 6.
FT	DOMAIN	999	1087	Fibronectin type-III 7.
FT	DOMAIN	1090	1179	Fibronectin type-III 8.
FT	DOMAIN	1203	1375	VMPA 3.
FT	DOMAIN	1388	1476	Fibronectin type-III 9.
FT	DOMAIN	1478	1567	Fibronectin type-III 10.
FT	DOMAIN	1569	1656	Fibronectin type-III 11.
FT	DOMAIN	1660	1747	Fibronectin type-III 12.
FT	DOMAIN	1758	1847	Fibronectin type-III 13.
FT	DOMAIN	1849	1937	Fibronectin type-III 14.
FT	DOMAIN	1939	2028	Fibronectin type-III 15.
FT	DOMAIN	2030	2119	Fibronectin type-III 16.
FT	DOMAIN	2121	2208	Fibronectin type-III 17.
FT	DOMAIN	2212	2298	Fibronectin type-III 18.
FT	DOMAIN	2329	2501	VMPA 4.
FT	DOMAIN	2525	2717	TSP N-terminal.
FT	DOMAIN	2456	2751	Nonhelical region (NC3).
FT	DOMAIN	2752	2899	Triple-helical region (COL2) with 1 imperfection.
FT	DOMAIN	2900	2942	Nonhelical region (NC2).
FT	DOMAIN	2943	3045	Triple-helical region (COL1) with 2 imperfections.
FT	DOMAIN	3046	3119	Nonhelical region (NC1).
FT	SITE	866	868	Cell attachment site (Potential).
FT	SITE	2784	2786	Cell attachment site (Potential).
FT	SITE	2896	2898	Cell attachment site (Potential).
FT	MOD_RES	2945	2945	Hydroxyproline (By similarity).
FT	MOD_RES	2948	2948	Hydroxyproline (By similarity).
FT	MOD_RES	2951	2951	Hydroxyproline (By similarity).
FT	MOD_RES	2960	2960	Hydroxyproline (By similarity).
FT	MOD_RES	2966	2966	Hydroxyproline (By similarity).
FT	MOD_RES	2969	2969	Hydroxyproline (By similarity).
FT	MOD_RES	2972	2972	Hydroxyproline (By similarity).
FT	MOD_RES	2984	2984	Hydroxyproline (By similarity).
FT	MOD_RES	3001	3001	Hydroxyproline (By similarity).
FT	MOD_RES	3004	3004	Hydroxyproline (By similarity).
FT	MOD_RES	3015	3015	Hydroxyproline (By similarity).
FT	MOD_RES	3024	3024	Hydroxyproline (By similarity).
FT	MOD_RES	3027	3027	Hydroxyproline (By similarity).
FT	MOD_RES	3030	3030	Hydroxyproline (By similarity).
FT	DOMAIN	869	872	Poly-Thr.
FT	CARBOHYD	704	704	O-linked (Xyl. . .) (Potential).
FT	CARBOHYD	802	802	O-linked (Xyl. . .) (chondroitin sulfate) (Potential).
FT	CARBOHYD	893	893	O-linked (Xyl. . .) (chondroitin sulfate) (Potential).
FT	CARBOHYD	985	985	O-linked (Xyl. . .) (chondroitin sulfate) (Potential).
FT	CARBOHYD	1769	1769	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	2212	2212	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	2533	2533	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	2684	2684	N-linked (GlcNAc. . .) (Potential).
FT	VARSPPLIC	25	1190	Missing (in isoform XIIB-1 and isoform XIIB-2).
FT	VARSPPLIC	3062	3064	/FTId=VSP_001150.
FT	VARSPPLIC	3065	3119	Epy -> GSG (in isoform XIIB-2 and isoform XIIB-2).
FT	VARSPPLIC	3119	3119	/FTId=VSP_001151.
FT	VARSPPLIC	3119	3119	Missing (in isoform XIIB-2 and isoform XIIB-2).
FT	VARSPPLIC	3119	3119	/FTId=VSP_001152.
FT	VARSPPLIC	3119	3119	MM, 981P959C6AB3251 CRC64,
FT	VARSPPLIC	3119	3119	Best local Similarity
FT	VARSPPLIC	3119	3119	Query Match
FT	VARSPPLIC	3119	3119	Score 134.5; DB 1; Length 3119;
FT	VARSPPLIC	3119	3119	Pred. No. 0.03;

Matches 56; Conservative 37; Mismatches 74; Indels 23; Gaps 11;

QY 3 DLVFLDKSGSV-LHNNIEIYFVEQLAKF-ISP-QLRMSFVPSRGTTLTKLTEDRE 59
 Db 444 DIVFLVGVSGYSIGTANFKVAFLEVLAKSPFISPNRVQSLVGVSDHPHTFETLKFNRR 503
 QY 60 QIRGQLEELQKVL-----PGDDTYMHGFEFASQIYENRQGRV--ASVITLTTGEL 112
 Db 504 ----VEDITIKAITFPFPGSGTNTGKAMTYVRKIKVFNPK-GSRSNVPMKMLITDGE- 556
 QY 113 HEDLFFYSERANSRDLAGIIVYCVGVDFNETQLARIAD--SKDHYFPVNDGQALQGI 170
 Db 557 SSDAF---RDPATLRLNSDVEIFAVGVKXDNVRSLEHAIASPPATHTYFTED-PDAQRRI 612
 QY 171 IHSILKKSCI 180
 Db 613 SFVLTSIGICL 622

RESULT 13
 ITAM_HUMAN STANDARD; PRT; 1152 AA.
 ID 01-JUL-1989 (Rel. 11, Created)
 AC P11215;
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1) (Neutrophil adherence receptor).
 DE Name:ITGAM; Synonyms:CD11B, CR3A;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP MEDLINE=88315033; Pubmed=2457584;
 RX Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;
 RT "The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b) alpha subunit. Cloning, primary structure, and relation to the integrins, von Willebrand factor and factor B.";
 RL J. Biol. Chem. 263:12403-12411 (1988).
 RN [2]
 RP MEDLINE=88190151; Pubmed=2833753;
 RX Arnaout M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;
 RT "Molecular cloning of the alpha subunit of human and guinea pig leukocyte adhesion glycoprotein Mo1: chromosomal localization and homology to the alpha subunits of integrins.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780 (1988).
 RN [3]
 RP MEDLINE=88257215; Pubmed=2454931; DOI=10.1083/jcb.106.6.2153;
 RX Arnaout M.A., Gupta S.K., Pierce M.W., Tenen D.G.;
 RT "Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor Mo1 (complement receptor type 3).";
 RL J. Cell Biol. 106:2153-2158 (1988).
 RN [4]
 RP MEDLINE=93123748; Pubmed=8419480;
 RX Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.F., Tenen D.G.;
 RT "Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-integrin gene family demonstrate remarkable conservation of genomic organization and suggest early diversification during evolution.";
 RL J. Immunol. 150:480-490 (1993).
 RN [5]
 RP MEDLINE=9098893; Pubmed=256162;
 RX Hickey D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L., Roth G.J.;
 RT "cDNA sequence for the alpha M subunit of the human neutrophil adherence receptor indicates homology to integrin alpha subunits.";

RL Proc. Natl. Acad. Sci. U.S.A. 86:257-261 (1989).
 RN [6]
 RP MEDLINE=92073318; Pubmed=1683702;
 RX Shelley C.S., Arnaout M.A.;
 RT "The promoter of the CD11b gene directs myeloid-specific and developmentally regulated expression.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:10525-10529 (1991).
 RN [7]
 RP MEDLINE=87076671; Pubmed=3539202; DOI=10.1016/0167-4838(86)90037-3;
 RX Pahl H.L., Rosmarin A.G., Tenen D.G.;
 RT "Characterization of the myeloid-specific CD11b promoter.";
 RL Blood 79:865-870 (1992).
 RN [8]
 RP MEDLINE=92144986; Pubmed=1346576;
 RX Pahl H.L., Rosmarin A.G., Tenen D.G.;
 RT "N-terminal sequence of human leukocyte glycoprotein Mo1: conservation across species and homology to platelet IIb/IIIa.";
 RL Biochim. Biophys. Acta 874:368-371 (1986).
 RN [9]
 RP MEDLINE=95171458; Pubmed=7867070; DOI=10.1016/0092-8674(95)90517-0;
 RX Lee J.O., Rieu P., Arnaout M.A., Liddington R.;
 RT "Crystal structure of the A domain from the alpha subunit of integrin CR3 (CD11b/CD18).";
 RL Cell 80:631-638 (1995).
 RN [10]
 RP MEDLINE=96363671; Pubmed=8747460; DOI=10.1016/S0969-2126(01)00271-4;
 RX Lee J.O., Bankston L.A., Arnaout M.A., Liddington R.C.;
 RT "Two conformations of the integrin A-domain (I-domain): a pathway for activation?";
 RL Structure 3:1333-1340 (1995).
 RN [11]
 RP MEDLINE=9832595; Pubmed=9687375; DOI=10.1016/S0969-2126(98)00093-8;
 RX Baldwin E.T., Saver R.W., Bryant G.L., Curry K.A., Fairbank M.B., Finzel B.C., Garlick R.L., Henthorn R.L., Horton N.C., Kelley L.L., Mildner A.M., Moon J.B., Mott J.B., Mutchler V.T., Tomich C.S., Watenpaugh K.D., Wiley V.H.;
 RT "Cation binding to the integrin CD11b I domain and activation model assessment.";
 RL Structure 6:923-935 (1998).
 RN [12]
 RP MEDLINE=98226734; Pubmed=9560195; DOI=10.1073/pnas.95.9.4870;
 RX O'Kvieg C., Springer T.A.;
 RT "Experimental support for a beta-propeller domain in integrin alpha-subunits and a calcium binding site on its lower surface.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875 (1998).
 CC -1- FUNCTION: Integrin alpha-M/beta-2 is implicated in various adhesive interactions of monocytes, macrophages and granulocytes as well as in mediating the uptake of complement-coated particles. It is identical with CR-3, the receptor for the IC3b fragment of the third complement component. It probably recognizes the R-G-D peptide in C3b. Integrin alpha-M/beta-2 is also a receptor for fibrinogen, factor X and ICAM1. It recognizes p1 and p2 peptides of fibrinogen gamma chain.
 CC -1- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-M associates with beta-2.
 CC -1- TISSUE SPECIFICITY: Type I membrane protein.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- GRANULOCYTES
 CC -1- DOMAIN: The integrin I-domain (inset) is a VMPA domain. Integrins with I-domains do not undergo protease cleavage.
 CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
 CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -1- SIMILARITY: Contains 1 VMPA domain.
 CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD11b entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11b.htm".

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.ebi.ac.uk/announcements/](http://www.ebi.ac.uk/announcements)
 CC or send an email to license@ebi.ac.uk).

DR EMBL/ J03925; AAA59544.1; -
 DR EMBL/ M18044; AAA59491.1; -
 DR EMBL/ J04145; AAA59903.1; -
 DR EMBL/ S52227; AAB24821.1; -
 DR EMBL/ S52152; AAB24821.1; JOINED.
 DR EMBL/ S52153; AAB24821.1; JOINED.
 DR EMBL/ S52154; AAB24821.1; JOINED.
 DR EMBL/ S52155; AAB24821.1; JOINED.
 DR EMBL/ S52157; AAB24821.1; JOINED.
 DR EMBL/ S52159; AAB24821.1; JOINED.
 DR EMBL/ S52161; AAB24821.1; JOINED.
 DR EMBL/ S52164; AAB24821.1; JOINED.
 DR EMBL/ S52165; AAB24821.1; JOINED.
 DR EMBL/ S52167; AAB24821.1; JOINED.
 DR EMBL/ S52169; AAB24821.1; JOINED.
 DR EMBL/ S52170; AAB24821.1; JOINED.
 DR EMBL/ S52173; AAB24821.1; JOINED.
 DR EMBL/ S52174; AAB24821.1; JOINED.
 DR EMBL/ S52180; AAB24821.1; JOINED.
 DR EMBL/ S52181; AAB24821.1; JOINED.
 DR EMBL/ S52184; AAB24821.1; JOINED.
 DR EMBL/ S52189; AAB24821.1; JOINED.
 DR EMBL/ S52191; AAB24821.1; JOINED.
 DR EMBL/ S52192; AAB24821.1; JOINED.
 DR EMBL/ S52203; AAB24821.1; JOINED.
 DR EMBL/ S52212; AAB24821.1; JOINED.
 DR EMBL/ S52213; AAB24821.1; JOINED.
 DR EMBL/ S52216; AAB24821.1; JOINED.
 DR EMBL/ S52219; AAB24821.1; JOINED.
 DR EMBL/ S52220; AAB24821.1; JOINED.
 DR EMBL/ S52221; AAB24821.1; JOINED.
 DR EMBL/ S52226; AAB24821.1; JOINED.
 DR EMBL/ M76724; AAA58410.1; -
 DR EMBL/ M84477; AAA51960.1; -
 DR PIR/ A31108; RWHU1B.
 DR PDB/ 1A8Y; Model; @-17-1152.
 DR PDB/ 1BHO; X-ray; 1/2=-.
 DR PDB/ 1BHQ; X-ray; 1/2=-.
 DR PDB/ 1IDN; X-ray; 1/2=-.
 DR PDB/ 1IDO; X-ray; @-140-331.
 DR PDB/ 1JLM; X-ray; @-143-334.
 DR PDB/ 1MLU; X-ray; A-137-331.
 DR PDB/ 1MF7; X-ray; A-144-337.
 DR PDB/ 1N92; X-ray; A-140-335.
 DR PDB/ 1NA5; X-ray; A-144-345.
 DR Genew; HGNC:6149; ITGAM.
 DR MIM; 120980; -
 DR CO; GO:0008305; C: Integrin complex; TAS.
 DR GO; GO:0007155; P: cell adhesion; TAS.
 DR Interpro; IPR000413; Integrin_alpha.
 DR Interpro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 3-
 DR Pfam; PF00357; Integrin_alpha; 1.
 DR Pfam; PF00092; VMA; 1.
 DR PRINTS; PRO1185; INTEGRIN.
 DR PRINTS; PRO0453; VWFADOMAIN.
 DR SMART; SM00191; Intc_alpha; 5.
 DR SMART; SM00327; VMA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PSS0233; VMAFA_1.
 DR 3D-structure; Calcium; Cell adhesion; Direct protein sequencing;
 KW Glycoprotein; Integrin; Magnesium; Receptor; Signal;
 KW Transmembrane.

FT SIGNAL 1 16 Integrin alpha-M.
 FT CHAIN 17 1152
 Query Match 14.1%; Score 132.5; DB 1; Length 1152;
 Best Local Similarity 26.3%; Pred. No. 0.014;
 Matches 52; Conservative 41; Mismatches 66; Indels 39; Gaps 11;

QY 3 DLVFIIDKSGSVL-HHNNIYYFVEQLAHKFIISPOLMSFIVSTRTGTMKLTED---- 57
 DQ 150 DIAFLIDGSGSIIIPHDRMKKEFVST-----VMEQLKSKTLFS-----LMQYSEPRRIH 199
 QY 58 -----RQIQGLEELQKVLPGEDTVMHSEGERASEQIYENRGQYTA-SVIALT 108
 DQ 200 FTTFKEFQNNPNPRLSLVPITQL--GRTHPATGIRKVRRLFNITNGARNAKPLIVIT 257
 QY 109 DGEIHEDLFYFSR--REANSRDLGAIYCGVGDENVETOLAR-----IADS--KDHVFP 159
 DQ 258 DGEKFGDPLGIEDYIPEADNR--GVIRYVIGGDARFSEKSRQELNTIASKPRDHFQ 314
 QY 160 VNDGFQALGGIHSILKK 177
 DQ 315 VNN-FBALKTITQNLREK 331

RESULT 14
 Q8CEK9 PRELIMINARY; PRT; 1182 AA.
 ID 08CEK9
 AC 08CEK9;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Mus musculus 2 days pregnant adult female ovary cDNA, RIKEN full-
 DE length enriched library, clone:E330019B14 product:hypothetical von
 DE Willebrand factor type A domain containing protein, full insert
 DE sequence.
 GN Name=E330026B02Rik;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Ovary;
 RX MEDLINE=9279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Ovary;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Ovary;
 RA The FANTOM Consortium;
 RT the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Ovary;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Ovary;
 RX MEDLINE=20530913; PubMed=11075861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Komno H., Akiyama Y., Nishi K., Kitanai T., Tashiro H., Itoh M.,
 RA Saito N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kasahara T.,
 RA Fujitake S., Inoue K., Togawa Y., Izawa M., Ohara E., Matsubara S., Kawai J.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Mizumoto M., Inoue Y., Kita A., Hayashizaki Y.;
 RT "RIKEN Integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Ovary;
 RA Adecchi U., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Hayashida K., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirokane T.,
 RA Hori F., Imocant K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Kato H., Kawai J., Kojima Y., Kondo S., Komno H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK054356; BAC3749.1; -.
 DR HSSP; P11215; IMF7.
 DR MGD; MGI:2444259; E330026B02R1K.
 DR InterPro; IPR02035; VWF_A.
 DR Pfam; PF00092; VMA; 6.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00327; VMA; 6.
 DR PROSITE; PS50234; VWF_A; 6.
 DR Hypothetical protein.
 KW SEQUENCE 1182 AA; 130946 MW; CA62P1926A426424 CRC64;
 SQ
 Query Match 13.7%; Score 129.5; DB 2; Length 1182;
 Best Local Similarity 22.0%; Pred. No. 0.026;
 Matches 42; Conservative 59; Mismatches 65; Indels 25; Gaps 9;
 QY 3 DLYEILDKSGSVL-HHNNETIYFPEOLAKHP-ISP-DLRMSFIYFSGRTGLMTL--D 57
 DB 435 DLYLILGSGSTOPTDHEMKTFPLSEVGMENLPHKRVGAVOYAUPTWDLFEFISKYSN 494
 QY 58 REQIQLQELQKVLPGSDT-----YMHGFERASEQIYENRQGYRTASVIALTDG 110
 DB 495 KPDLGKAIENIRQW--GGNTVTGAALNFTLKLGRKKE-----RSKYPCHLVLTNG 546
 QY 111 ELHEDLFYFSERANRSDGAIYCYGVQDFNETQLARLADSKDHFVPVNDGQALQGI 170
 DB 547 MSRSRSV-----GPAHKLRERINIRVHAIGVKEANOTQRLRAGEKRYVYVHE-FDALRNI 601
 QY 171 IHSILKSCIE 181
 DB 602 RNOVVDETCAE 612
 RESULT 15
 CALC_CHICK STANDARD; PRT; 3124 AA.
 AC P13944; Q04509;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Collagen alpha 1(XII) chain precursor (Fibronchimerin).
 GN Name=COL12A1;
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 NC NCBI_TaxID=9031;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=White Leghorn;
 RX MEDLINE=92011862; PubMed=1918137; DOI=10.1083/jcb.115.1.209;
 RA Yamagata M., Yamada K.M., Yamada S.S., Shinomura T., Tanaka H.,
 RA Nishida Y., Ohara M., Kimita K.;
 RT "The complete primary structure of type XII collagen shows a chimeric
 RT molecule with reiterated fibronectin type III motifs, von Willebrand
 RT factor A motifs, a domain homologous to a noncollagenous region of
 RT type IX collagen, and short collagenous domains with an Arg-Gly-Asp
 RT site.";
 RL J. Cell Biol. 115:209-221(1991).
 RN [2]
 RP SEQUENCE OF 2456-3124 FROM N.A., AND SEQUENCE OF 2772-2794 AND
 RP 2846-2873.
 RX MEDLINE=90062079; PubMed=2584192;
 RA Gordon M.K., Gerecke D.R., Dublet B., van der Rest M., Olsen B.R.;
 RT "Type XII collagen. A large multidomain molecule with partial homology
 RT to type IX collagen.";
 RL J. Biol. Chem. 264:19772-19778(1989).
 RN [3]
 RP SEQUENCE OF 2960-3076 FROM N.A.
 RX MEDLINE=87317590; PubMed=3476925;
 RA Gordon M.K., Gerecke D.R., Olsen B.R.;
 RT "Type XII collagen: distinct extracellular matrix component discovered
 RT by cDNA cloning.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:6040-6044(1987).
 RN [4]
 RP SEQUENCE OF 1-1283 FROM N.A. (ISOFORM SHORT), AND ALTERNATIVE
 RP SPLICING.
 RC TISSUE=Embryo;
 RX MEDLINE=93042014; PubMed=1420368; DOI=10.1016/0167-4781(92)90145-P;
 RA Tneub J., Tneub B.;
 RT "The two splice variants of collagen XII share a common 5' end.";
 RL Biochim. Biophys. Acta 1171:97-98(1992).
 RN [5]
 RP ALTERNATIVE SPLICING.
 RX MEDLINE=95370352; PubMed=7642694; DOI=10.1083/jcb.130.4.1005;
 RA Koch M., Bornmann B., Matchison M., Hagios C., Tneub B., Chiquet M.;
 RT "Large and small splice variants of collagen XII: differential
 RT expression and ligand binding.";
 RL J. Cell Biol. 130:1005-1014(1995).
 CC -1- FUNCTION: Type XII collagen interacts with type I collagen-
 CC containing fibrils, the COL1 domain could be associated with the
 CC surface of the fibrils, and the COL2 and NC3 domains may be
 CC localized in the perifibrillar matrix.
 CC -1- SUBUNIT: Trimer of identical chains each containing 190 kDa of
 CC nontriple-helical sequences.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Comment=The final tissue form of collagen XII may contain
 CC homotrimers of either isoform long or isoform short or any
 CC combination of isoform long and isoform short. Only isoform long
 CC is a proteoglycan. Isoform long has more restricted expression
 CC in embryonic tissue than isoform short;
 CC Name=Long;
 CC IsoId=P13944-1; Sequence=Displayed;
 CC Name=Short;
 CC IsoId=P13944-2; Sequence=VSP_001148;
 CC -1- TISSUE SPECIFICITY: Type XII collagen is present in tendons,
 CC ligaments, perichondrium, and perosteum, all dense connective
 CC tissues containing type I collagen.
 CC -1- DOMAIN: This sequence defines five distinct domains, two triple-
 CC helical domains (COL1 and COL2) and three nontriple-helical
 CC domains (NC1, NC2, and NC3).
 CC -1- PTM: The triple-helical tail is stabilized by disulfide bonds at
 CC each end.
 CC -1- PTM: Prolines at the third position of the tripeptide repeating
 CC unit (G-X-Y) are hydroxylated in some or all of the chains.
 CC -1- PTM: O-glycosylated; glycosaminoglycan of chondroitin-sulfate type
 CC (By similarity).
 CC -1- SIMILARITY: Belongs to the fibril-associated collagens with
 CC interrupted helices (FACIT) family.

CC -1- SIMILARITY: Contains 18 fibronectin type III domains.
 CC -1- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
 CC -1- SIMILARITY: Contains 4 WFA domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).

DR EMBL; D00824; BAA00701.1; -;
 DR EMBL; X61024; CAA43358.1; -;
 DR EMBL; M17375; AAA48718.1; -;
 DR EMBL; J05137; AAA46635.1; -;
 DR EMBL; X67327; CAA47744.1; -;
 DR PIR; A40020; A40020.
 DR HSP; P56199; I0C5.
 DR InterPro; IPR008160; Collagen.
 DR InterPro; IPR008985; CNA_1like_1ec_g1.
 DR InterPro; IPR003961; FN_III.
 DR InterPro; IPR008957; FN_III-like.
 DR InterPro; IPR003129; TSP_N.
 DR InterPro; IPR002035; WFA_2.
 DR Pfam; PF00041; fn3; 17.
 DR Pfam; PF02210; TSP_N; 1.
 DR Pfam; PF00092; WFA_4.
 DR PRINTS; PR00453; WFAADOMAIN.
 DR SMART; SM00060; FN3; 18.
 DR SMART; SM00327; WFA; 4.
 DR PROSITE; PS50853; FN3; 18.
 DR PROSITE; PS50234; WFA; 4.
 DR Alternative splicing; Cell adhesion; Collagen;
 DR Direct protein sequencing; Extracellular matrix; Glycoprotein;
 DR Hydroxylation; Repeat; Signal; Structural protein.
 FT SIGNAL 1 24 Potential.
 FT CHAIN 25 3124 Collagen alpha 1(XIII) chain.
 FT DOMAIN 25 3112 Fibronectin type-III 1.
 FT DOMAIN 139 311 WFA 1.
 FT DOMAIN 332 421 Fibronectin type-III 2.
 FT DOMAIN 439 615 WFA 2.
 FT DOMAIN 630 718 Fibronectin type-III 3.
 FT DOMAIN 721 809 Fibronectin type-III 4.
 FT DOMAIN 812 902 Fibronectin type-III 5.
 FT DOMAIN 905 993 Fibronectin type-III 6.
 FT DOMAIN 995 1083 Fibronectin type-III 7.
 FT DOMAIN 1086 1175 Fibronectin type-III 8.
 FT DOMAIN 1199 1371 WFA 3.
 FT DOMAIN 1386 1472 Fibronectin type-III 9.
 FT DOMAIN 1474 1564 Fibronectin type-III 10.
 FT DOMAIN 1566 1654 Fibronectin type-III 11.
 FT DOMAIN 1655 1745 Fibronectin type-III 12.
 FT DOMAIN 1756 1845 Fibronectin type-III 13.
 FT DOMAIN 1847 1935 Fibronectin type-III 14.
 FT DOMAIN 1937 2026 Fibronectin type-III 15.
 FT DOMAIN 2028 2117 Fibronectin type-III 16.
 FT DOMAIN 2119 2206 Fibronectin type-III 17.
 FT DOMAIN 2210 2294 Fibronectin type-III 18.
 FT DOMAIN 2327 2500 TSP N-terminal.
 FT DOMAIN 2524 2716 Nonhelical region (NC2).
 FT DOMAIN 2455 2750 Triple-helical region (COL2) with 1
 FT DOMAIN 2751 2902 imperfection.
 FT DOMAIN 2903 2945 Nonhelical region (NC2).
 FT DOMAIN 2946 3048 Triple-helical region (COL1) with 2
 FT DOMAIN 3049 3124 imperfections.
 FT DOMAIN 3086 3096 Nonhelical region (NC1).
 FT DOMAIN 3111 3123 Asp/Glu-rich (acidic).
 FT SITE 2899 2901 Arg/Lys-rich (basic).
 FT 2899 2901 Cell attachment site (Potential).

FT CARBOHYD	32	32	N-linked (GlcNAc. .) (Potential).
FT CARBOHYD	797	797	O-linked (Xyl. .) (chondroitin sulfate)
FT CARBOHYD	890	890	O-linked (Xyl. .) (chondroitin sulfate)
FT CARBOHYD	981	981	O-linked (Xyl. .) (chondroitin sulfate)
FT CARBOHYD	1006	1006	N-linked (GlcNAc. .) (Potential).
FT CARBOHYD	1032	1032	N-linked (GlcNAc. .) (Potential).
FT CARBOHYD	1044	1044	N-linked (GlcNAc. .) (Potential).
FT CARBOHYD	1512	1512	N-linked (GlcNAc. .) (Potential).
FT CARBOHYD	1767	1767	N-linked (GlcNAc. .) (Potential).
FT CARBOHYD	2210	2210	N-linked (GlcNAc. .) (Potential).
FT CARBOHYD	2273	2273	N-linked (GlcNAc. .) (Potential).
FT CARBOHYD	2532	2532	N-linked (GlcNAc. .) (Potential).
FT CARBOHYD	2683	2683	N-linked (GlcNAc. .) (Potential).
FT VARSPLIC	25	1188	Missing (in isoform Shortt).
FT CONFLICT	1258	1258	/FTid=VSP_001148.
FT CONFLICT	1264	1264	T -> S (in Ref. 4).
FT CONFLICT	2759	2759	D -> E (in Ref. 4).
FT CONFLICT	2803	2803	P -> A (in Ref. 2).
FT CONFLICT	2977	2977	L -> F (in Ref. 2).
FT CONFLICT	3075	3076	V -> F (in Ref. 2).
FT CONFLICT	3124	340578	QP -> AG (in Ref. 3).
FT SEQUENCE	3124	AA; 340578	MM; 094285AFETf346CF CRC64;

Query Match 13.6%; Score 128; DB 1; Length 3124;
 Best Local Similarity 26.3%; Pred. No. 0.11; Indels 16; Gaps 9;
 Matches 49; Conservative 42; Mismatches 79;

QY	3	DLYFIIDKSGSV-LHMNEIYYFVEQLAHKF--ISPO-LRMSFIVFSTRTGLMKLT--E 56
DB	2327	DIYFLTDASWSISIGDNNKVKVFENTVGAFLDINPAGIQVSLVQYSDAQSEFKLTWFD 2386
QY	57	DREQIRGELBELQKVLPGCDTYHMEGRASQIY-YENRQGRTAIVIALTNGELHED 115
DB	2387	DKQALGALGNVQ--YRGWTRGKALFTFKELTWESGMRRCVPRVLVVTGSRQDE 2444
QY	116	LFYSEBRANRSRLAIVCVGVKDPNETOLARIAD--SKDHYFPVNDGFQALOGIHS 173
DB	2445	V---RKAAIVYIQHSGSVFVGVADVYNELAKIAPSRHVFYVD--FDAFEKIQDN 2499
QY	174	ILKKSQ 179
DB	2500	LVTFPVC 2505

Search completed: June 13, 2005, 20:03:17
 Job time : 61.7608 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 19:37:17 ; Search time 64.0514 Seconds
(without alignments)
1044.623 Million cell updates/sec

Title: us-09-970-076-2_COPY_44_216
Perfect score: 1 DLYFILDKSSVLIHMEIY.....HVEPVNDGQALQGIHSIL 173
Sequence:

Scoring table: BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A_GeneSeq_16Dec04:*

1: geneSeq1980s:*\n2: geneSeq1990s:*\n3: geneSeq2000s:*\n4: geneSeq2001s:*\n5: geneSeq2002s:*\n6: geneSeq2003as:*\n7: geneSeq2003bs:*\n8: geneSeq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	899	100.0	297	4	AAM38976 Human pol
2	899	100.0	328	7	Adi00558 Human TAN
3	899	100.0	328	7	Adm64584 Human TAN
4	899	100.0	333	3	AA801422 Human TAN
5	899	100.0	333	5	ABP54905 Human ant
6	899	100.0	333	7	Adi00534 Human TAN
7	899	100.0	333	7	Adm64568 Human von
8	899	100.0	342	7	Adi00554 Human TAN
9	899	100.0	342	7	Adm64580 Human TAN
10	899	100.0	345	7	Adi00556 Human TAN
11	899	100.0	345	7	Adm64582 Human TAN
12	899	100.0	368	5	ABP54903 Human ant
13	899	100.0	384	5	Adm64586 TANGO197
14	899	100.0	403	4	AAE01439 Human gen
15	899	100.0	403	5	ABG63874 Human alb
16	899	100.0	403	8	Adi77139 Albumin f
17	899	100.0	460	7	Adi00562 Human TAN
18	899	100.0	460	7	Adi00560 Human TAN
19	899	100.0	479	7	Adi00564 Human TAN
20	899	100.0	504	7	Adi00566 Human TAN
21	899	100.0	529	7	Adi00568 Human TAN
22	899	100.0	538	7	ABE97421 Human chi
23	899	100.0	540	7	Adi00544 Human TAN
24	899	100.0	540	7	Adm64588 TANGO197
25	899	100.0	549	7	Adi00546 Human TAN

26	899	100.0	549	7	Adi00542 Human TAN
27	899	100.0	549	7	Adi00548 Human TAN
28	899	100.0	549	7	Adm64590 TANGO197
29	899	100.0	549	7	Adm64592 TANGO197
30	899	100.0	551	7	Adi00550 Human TAN
31	899	100.0	551	7	Adm64576 Mouse TAN
32	899	100.0	564	5	ABB90750 Human Tum
33	899	100.0	564	5	ABB90724 Human Tum
34	899	100.0	564	5	ABP54904 Human ant
35	899	100.0	564	6	ABU54457 Human tum
36	899	100.0	564	6	ABU54431 Human tum
37	899	100.0	564	7	Adi00552 Human TAN
38	899	100.0	564	7	ADJ70017 Human hea
39	899	100.0	564	7	Adm64578 Human TAN
40	899	100.0	564	8	ADR48216 Human tum
41	894	99.4	403	4	AAE01469 Human gen
42	894	99.4	403	5	ABG63873 Human alb
43	894	99.4	403	8	Adi77138 Albumin f
44	890	99.0	562	5	ABB90731 Mouse Tum
45	890	99.0	562	5	ABB90785 Mouse Tum

ALIGNMENTS

RESULT 1	
AA83976	
ID	AA83976 standard; protein; 297 AA.
XX	
AC	AA83976;
XX	
DT	22-OCT-2001 (first entry)
XX	
DE	Human polypeptide SEQ ID NO 2121.
XX	
OS	Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia.
XX	
OS	Homo sapiens.
XX	
XX	
PN	MO20015312-A1.
XX	
PD	26-JUL-2001.
XX	
PF	26-DEC-2000; 2000WO-US034263.
XX	
PR	23-DEC-1999; 99US-00471275.
XX	
PR	21-JAN-2000; 2000US-00488725.
XX	
PR	25-APR-2000; 2000US-00552317.
XX	
PR	20-JUN-2000; 2000US-00598042.
XX	
PR	19-JUL-2000; 2000US-00620312.
XX	
PR	03-AUG-2000; 2000US-00653450.
XX	
PR	14-SEP-2000; 2000US-00662191.
XX	
PR	19-OCT-2000; 2000US-00693036.
XX	
PR	29-NOV-2000; 2000US-00727344.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PI	Tang YT, Liu C, Aeundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D; Wang Z, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA; Zhou P, Goodrich R, Drmanac RT;
PI	
XX	WPI; 2001-442253/47.
DR	N-PSDB; AAI58132.
XX	
PT	Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries.
XX	
PS	Example 4; SEQ ID NO 2121; 10078bp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and the
CC encoded polypeptides (AA38642-AA44213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC activation/inhibition activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX

SQ Sequence 297 AA;

Query Match 100.0%; Score 899; DB 4; Length 297;
Best Local Similarity 100.0%; Pred. No. 2.4e-92;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Dy 1 DLVFIIDKSGSVLHHNNEIYFVEQLAKRFISPOLRMSFIVFSTRGTTLMKLTEDRQIR 60
Db 44 DLVFIIDKSGSVLHHNNEIYFVEQLAKRFISPOLRMSFIVFSTRGTTLMKLTEDRQIR 103
Qy 61 QGLEELQKVLPGGDTYMHGEFERASEQIYYENRQGYRTASVITALTDEGLHEDLFFYSER 120
Db 104 QGLEELQKVLPGGDTYMHGEFERASEQIYYENRQGYRTASVITALTDEGLHEDLFFYSER 163
Qy 121 EANRSRLGAIIVYCVGKDFNETQLARIADSKDHVPVNDGFQALQGIHSIL 173
Db 164 EANRSRLGAIIVYCVGKDFNETQLARIADSKDHVPVNDGFQALQGIHSIL 216

RESULT 2

AD100558 ID AD100558 standard; protein; 328 AA.

XX AD100558;

DT 22-Apr-2004 (first entry)

XX Human TANGO 197 Histag fusion protein - plasmid pO615.

XX fusion; von Willebrand factor A-like domain; vWF; antibacterial;
KM cutaneous; inhalation anthrax; human; TANGO 197 Histag fusion; mutant;
KM plasmid pO615; mutcin.

XX Homo sapiens.

OS Synthetic.

XX US2003144193-A1.

PD 31-JUL-2003.

XX 24-JUL-2002; 2002US-00201292.

XX 20-DEC-2001; 2001US-00038307.

XX (ROTT/) ROTTMAN J B.

XX (OKEE/) O'KEEFE T L.

XX (OZKA/) OZKAYNAK E.

XX (HEAL/) HEALEY J J.

XX Rottman JB, O'Keefe TL, Ozkaynak E, Healey JJ;

XX WPI; 2003-720708/68.

XX N-PSDB; AD100557.

XX New TANGO 197 and/or TANGO 216 fusion polypeptides for preventing or
PT ameliorating symptoms of anthrax comprises a von Willebrand factor A-like

PT domain (vWF) amino acid sequence and an amino acid sequence heterologous
PT to the vWF.

XX Claim 45; SEQ ID NO 26; 86pp; English.

XX The invention relates to a novel fusion polypeptide comprising a von
CC Willebrand factor A-like domain (vWF) amino acid sequence and an amino
CC acid sequence heterologous to the vWF. The polypeptide of the invention
CC demonstrates antibacterial activities whilst the composition and method
CC may be useful in preventing or ameliorating the symptoms of cutaneous
CC and/or inhalation anthrax. The current sequence is that of the human
CC TANGO 197 Histag fusion protein of the invention.
XX

SQ Sequence 328 AA;

Query Match 100.0%; Score 899; DB 7; Length 328;
Best Local Similarity 100.0%; Pred. No. 2.7e-92;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Dy 1 DLVFIIDKSGSVLHHNNEIYFVEQLAKRFISPOLRMSFIVFSTRGTTLMKLTEDRQIR 60
Db 44 DLVFIIDKSGSVLHHNNEIYFVEQLAKRFISPOLRMSFIVFSTRGTTLMKLTEDRQIR 103
Qy 61 QGLEELQKVLPGGDTYMHGEFERASEQIYYENRQGYRTASVITALTDEGLHEDLFFYSER 120
Db 104 QGLEELQKVLPGGDTYMHGEFERASEQIYYENRQGYRTASVITALTDEGLHEDLFFYSER 163
Qy 121 EANRSRLGAIIVYCVGKDFNETQLARIADSKDHVPVNDGFQALQGIHSIL 173
Db 164 EANRSRLGAIIVYCVGKDFNETQLARIADSKDHVPVNDGFQALQGIHSIL 216

RESULT 3

ADM64584 ID ADM64584 standard; protein; 328 AA.

XX ADM64584;

DT 03-JUN-2004 (first entry)

XX Human TANGO197-His tag fusion protein #2.

XX antibacterial; gene therapy;
KM von Willebrand factor A-like domain amino acid sequence;
KM vWF amino acid sequence; anthrax; Bacillus anthracis; cutaneous anthrax;
KM inhalation anthrax; human; TANGO197; his tag; fusion protein.

XX Homo sapiens.

OS Synthetic.

XX US2003134786-A1.

PD 17-JUL-2003.

XX 20-DEC-2001; 2001US-00038307.

XX 20-DEC-2001; 2001US-00038307.

XX (ROTT/) ROTTMAN J B.

XX (OKEE/) O'KEEFE T L.

XX (OZKA/) OZKAYNAK E.

XX (HEAL/) HEALEY J J.

XX Rottman JB, O'Keefe TL, Ozkaynak E, Healey JJ;

XX WPI; 2003-829643/77.

XX N-PSDB; ADM64583.

XX New TANGO 197 and/or TANGO 216 fusion polypeptides for preventing or
PT ameliorating symptoms of anthrax comprises a von Willebrand factor A-like
PT domain (vWF) amino acid sequence and an amino acid sequence heterologous
PT to the vWF.

PS Claim 44; SEQ ID NO 26; 64pp; English.

XX The invention describes a fusion polypeptide comprising a von Willebrand
 CC factor A-like domain (WVF) amino acid sequence and an amino acid sequence
 CC heterologous to the WVF. Also described are: a method of preventing or
 CC ameliorating a symptom of anthrax in a subject thought to be at risk for
 CC exposure to or suspected of having been exposed to Bacillus anthracis;
 CC and a pharmaceutical composition comprising the novel fusion polypeptide.
 CC The composition and method are useful in preventing or ameliorating
 CC symptoms of cutaneous and/or inhalation anthrax. This is the amino acid
 CC sequence of a fusion protein comprising mature human TANGO197, thrombin
 CC cleavage site and his tag that can be used to treat exposure to or
 CC prevent a symptom of anthrax.

XX
 SQ Sequence 328 AA;

Query Match 100.0%; Score 899; DB 7; Length 328;
 Best Local Similarity 100.0%; Pred. No. 2,7e-92;
 Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLYFILDKSGSVLHMHNEIYFVEQLAHKFIISPOLRMSFIVFSTRGTTMLKLTEDREQIR 60
 DB 44 DLYFILDKSGSVLHMHNEIYFVEQLAHKFIISPOLRMSFIVFSTRGTTMLKLTEDREQIR 103

QY 61 QGLEELQKVLPGDPTVMHEGFERASEQIYYENRQGYRTASVITALTGELHEDLFFYSER 120
 DB 104 QGLEELQKVLPGDPTVMHEGFERASEQIYYENRQGYRTASVITALTGELHEDLFFYSER 163

QY 121 EANSRDLGAIIVCVGVKDFNETQLARIADSKDVFPVNDGFQALQGIHSIL 173
 DB 164 EANSRDLGAIIVCVGVKDFNETQLARIADSKDVFPVNDGFQALQGIHSIL 216

RESULT 4
 ID AAB01422 standard; protein; 333 AA.

XX
 AC AAB01422;
 DT 20-OCT-2000 (first entry)
 DE Human TANGO 197.

XX TANGO; 128; 140; 197; 212; 213; 224; 239; modulating agent; asthma;
 KW graft versus-host diseases; rheumatoid arthritis; psoriasis;
 KW inflammatory bowel diseases; septic shock; ulcerative colitis;
 KW Crohn's disease; chronic myelogenous leukemia; cancer; liver disease;
 KW Hodgkin's disease; osteoarthritis; Lyme's disease; cachexia;
 KW autoimmune disease; myasthenia gravis; autoimmune diabetes;
 KW systemic lupus erythematosus; transgenic animal; diagnosis; prognosis;
 KW prophylactic; therapeutic; human.

XX
 OS Homo sapiens.
 XX
 PN WO200039284-A1.
 PD 06-JUL-2000.
 PF 23-DEC-1999; 99WO-US031025.
 XX
 PR 30-DEC-1998; 98US-00223546.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Holtzman DA;
 XX
 DR WPI; 2000-465743/40.
 DR N-PSDB; AAA47455.
 XX
 PT Novel nucleic acid sequences encoding TANGO-128, 140, 197, 212, 213, 224
 PT and 239 polypeptides useful for the treatment of asthma, rheumatoid
 PT arthritis, psoriasis and autoimmune diseases.
 XX

PS Claim 8; Fig 4; 209pp; English.

XX Nucleic acids encoding TANGO polypeptides are useful as modulating agents
 CC for regulating cellular processes like asthma, graft versus-host
 CC diseases, rheumatoid arthritis, psoriasis, inflammatory bowel disease,
 CC septic shock, ulcerative colitis, Crohn's disease, chronic myelogenous
 CC leukemia, cancer, liver disease, Hodgkin's disease, osteoarthritis,
 CC Lyme's disease, cachexia and autoimmune diseases e.g. myasthenia gravis,
 CC autoimmune diabetes and systemic lupus erythematosus. The nucleic acids
 CC are also useful for producing transgenic animals and the TANGO
 CC polypeptides themselves. Partial TANGO-128, 140, 197, 212, 213, 224, 239
 CC sequences are useful in forensic biology, for diagnostic assays,
 CC prognostic assays, pharmacogenomics and for monitoring clinical trials.
 CC TANGO polypeptides are suitable for both prophylactic and therapeutic
 CC methods for treating a subject at risk of a disorder or having a disorder
 CC associated with aberrant TANGO expression. A wide range of cellular
 CC disorders can be treated

XX
 SQ Sequence 333 AA;

Query Match 100.0%; Score 899; DB 3; Length 333;
 Best Local Similarity 100.0%; Pred. No. 2,8e-92;
 Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLYFILDKSGSVLHMHNEIYFVEQLAHKFIISPOLRMSFIVFSTRGTTMLKLTEDREQIR 60
 DB 44 DLYFILDKSGSVLHMHNEIYFVEQLAHKFIISPOLRMSFIVFSTRGTTMLKLTEDREQIR 103

QY 61 QGLEELQKVLPGDPTVMHEGFERASEQIYYENRQGYRTASVITALTGELHEDLFFYSER 120
 DB 104 QGLEELQKVLPGDPTVMHEGFERASEQIYYENRQGYRTASVITALTGELHEDLFFYSER 163

QY 121 EANSRDLGAIIVCVGVKDFNETQLARIADSKDVFPVNDGFQALQGIHSIL 173
 DB 164 EANSRDLGAIIVCVGVKDFNETQLARIADSKDVFPVNDGFQALQGIHSIL 216

RESULT 5
 ID ABP54905 standard; protein; 333 AA.

XX
 AC ABP54905;
 DT 08-JAN-2003 (first entry)
 DE Human anthrax toxin receptor.

XX Anthrax; toxin; receptor; human; antibacterial.
 KW
 KW Homo sapiens.

XX
 OS Homo sapiens.
 XX
 FH Key
 FT Peptide 1..27
 FT Protein /label= Signal_peptide
 FT /label= Mature_protein
 XX
 PN WO200246228-A2.
 PD 13-JUN-2002.
 PF 03-OCT-2001; 2001WO-US030941.
 XX
 PR 05-DEC-2000; 2000US-0251481P.
 XX
 PA (WISC) WISCONSIN ALUMNI RES FOUND.
 XX
 PI Young JAT, Bradley KA, Collier RJ, Mogridge JS;
 XX
 DR WPI; 2002-713235/77.
 DR N-PSDB; ABV73883.
 XX
 PT Novel isolated polypeptide useful for identifying agent that prevents or

PT reduces effect of anthrax toxin on host cell, for treating human or non-
PT human animal suffering from anthrax.
XX
PS Claim 1, Page 42-43; 45pp; English.
XX
CC The present sequence is the protein sequence of a polypeptide identified
CC as a human anthrax toxin receptor (ATR) polypeptide on the basis of
CC identity to a newly isolated human ATR (see ABP54903). The 2 polypeptides
CC are identical between amino acids 1-317, but differ thereafter at the C-
CC terminus. The present polypeptide, encoded by IMAGE CLONE 4563020, has no
CC previously known function, and there has been no prior indication that it
CC is a complete or partial ATR. The invention provides ATR polypeptides and
CC polynucleotides, vectors, host cells, and transgenic and knock-out
CC animals. It also provides methods for identifying molecules that bind the
CC ATR and which reduce the toxicity of anthrax toxin. A claimed method for
CC treating anthrax in a human or animal involves administering an agent
CC that inhibits binding between anthrax toxin protective antigen (PA) and
CC ATR at a level effective to reduce the severity of anthrax. Suitable
CC agents include the present polypeptide or a PA-binding fragment of it, a
CC PA-binding polypeptide at least 80% identical to these, a fusion protein,
CC a monoclonal or polyclonal antibody, a polysaccharide, a lipid or a
CC nucleic acid
XX
SQ Sequence 333 AA;
XX
Query Match 100.0%; Score 899; DB 5; Length 333;
Best Local Similarity 100.0%; Pred. No. 2.8e-92;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DLVFIIDKSGSVLHHNNEIYFVEQLAHKFIISPOLRMSFIVFSTRGTTLMKLTEDREQIR 60
DB 44 DLVFIIDKSGSVLHHNNEIYFVEQLAHKFIISPOLRMSFIVFSTRGTTLMKLTEDREQIR 103
QY 61 QGSEELQKVLPGGDTYMHGFEFERSAQIYYENRGYRTASVIALTDGELHEDLFFYSER 120
DB 104 QGSEELQKVLPGGDTYMHGFEFERSAQIYYENRGYRTASVIALTDGELHEDLFFYSER 163
QY 121 EANRSRLGAIYVCVGVKDFNETOLARIADSKDHVPVNDGFQALQGIHSIL 173
DB 164 EANRSRLGAIYVCVGVKDFNETOLARIADSKDHVPVNDGFQALQGIHSIL 216
RESULT 6
ADI00534
ID ADI00534 standard; protein; 333 AA.
XX
AC ADI00534;
XX
DT 22-APR-2004 (first entry)
XX
DE Human TANGO 197 protein.
XX
KM fusion; von Willebrand factor A-like domain; vWF; antibacterial;
XX cutaneous; inhalation anthrax; human; TANGO 197.
XX
OS Homo sapiens.
XX
PN US2003144193-A1.
XX
PD 31-JUL-2003.
XX
PF 24-JUL-2002; 2002US-00201292.
XX
PR 20-DEC-2001; 2001US-00038307.
XX
PA (ROTT/) ROTTMAN J B.
PA (OKEE/) O'KEEFE T L.
PA (OZKA/) OZKAYNAK E.
PA (HEAL/) HEALEY J J.
XX
PI Rottman JB, O'keefe TL, Ozkaynak E, Healey JJ;
XX
DR WPI; 2003-720708/68.

DR N-PSDB; ADI00533.
XX
PT New TANGO 197 and/or TANGO 216 fusion polypeptides for preventing or
PT ameliorating symptoms of anthrax comprises a von Willebrand factor A-like
PT domain (vWF) amino acid sequence and an amino acid sequence heterologous
PT to the vWF.
XX
PS Claim 26; SEQ ID NO 2; 86pp; English.
XX
CC The invention relates to a novel fusion polypeptide comprising a von
CC Willebrand factor A-like domain (vWF) amino acid sequence and an amino
CC acid sequence heterologous to the vWF. The polypeptide of the invention
CC demonstrates antibacterial activities whilst the composition and method
CC may be useful in preventing or ameliorating the symptoms of cutaneous
CC and/or inhalation anthrax. The current sequence is that of the human
CC TANGO 197 protein of the invention.
XX
SQ Sequence 333 AA;
XX
Query Match 100.0%; Score 899; DB 7; Length 333;
Best Local Similarity 100.0%; Pred. No. 2.8e-92;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DLVFIIDKSGSVLHHNNEIYFVEQLAHKFIISPOLRMSFIVFSTRGTTLMKLTEDREQIR 60
DB 44 DLVFIIDKSGSVLHHNNEIYFVEQLAHKFIISPOLRMSFIVFSTRGTTLMKLTEDREQIR 103
QY 61 QGSEELQKVLPGGDTYMHGFEFERSAQIYYENRGYRTASVIALTDGELHEDLFFYSER 120
DB 104 QGSEELQKVLPGGDTYMHGFEFERSAQIYYENRGYRTASVIALTDGELHEDLFFYSER 163
QY 121 EANRSRLGAIYVCVGVKDFNETOLARIADSKDHVPVNDGFQALQGIHSIL 173
DB 164 EANRSRLGAIYVCVGVKDFNETOLARIADSKDHVPVNDGFQALQGIHSIL 216
RESULT 7
ADM64568
ID ADM64568 standard; protein; 333 AA.
XX
AC ADM64568;
XX
DT 03-JUN-2004 (first entry)
XX
DE Human von Willebrand factor A-like domain protein TANGO197.
XX
KM antibacterial; gene therapy;
XX von Willebrand factor A-like domain amino acid sequence;
XX vWF amino acid sequence; anthrax; Bacillus anthracis; cutaneous anthrax;
XX inhalation anthrax; human; TANGO197.
XX
OS Homo sapiens.
XX
PN US2003134786-A1.
XX
PD 17-JUL-2003.
XX
PF 20-DEC-2001; 2001US-00038307.
XX
PR 20-DEC-2001; 2001US-00038307.
XX
PA (ROTT/) ROTTMAN J B.
PA (OKEE/) O'KEEFE T L.
PA (OZKA/) OZKAYNAK E.
PA (HEAL/) HEALEY J J.
XX
PI Rottman JB, O'keefe TL, Ozkaynak E, Healey JJ;
XX
DR WPI; 2003-829643/77.
XX
DR N-PSDB; ADM64567.
XX
PT New TANGO 197 and/or TANGO 216 fusion polypeptides for preventing or
PT ameliorating symptoms of anthrax comprises a von Willebrand factor A-like

PT domain (vWF) amino acid sequence and an amino acid sequence heterologous
PT to the vWF.
PS Claim 26; SEQ ID NO 2; 64pp; English.
XX
CC The invention describes a fusion polypeptide comprising a von Willebrand
CC factor A-like domain (vWF) amino acid sequence and an amino acid sequence
CC heterologous to the vWF. Also described are: a method of preventing or
CC ameliorating a symptom of anthrax in a subject thought to be at risk for
CC exposure to or suspected of having been exposed to *Bacillus anthracis*;
CC and a pharmaceutical composition comprising the novel fusion polypeptide.
CC The composition and method are useful in preventing or ameliorating
CC symptoms of cutaneous and/or inhalation anthrax. This is the amino acid
CC sequence of a human von Willebrand factor A-like domain (vWF) amino acid
CC sequence TANGO197.
XX
SQ Sequence 333 AA;
SQ
Query Match 100.0%; Score 899; DB 7; Length 333;
Best Local Similarity 100.0%; Pred. No. 2.8e-92;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DLVPIIDKSGSVLHNMNIYFVQOLAHKFTSPQLRMSFTVFSRGTTLTKLTEDRQIR 60
DB 44 DLVPIIDKSGSVLHNMNIYFVQOLAHKFTSPQLRMSFTVFSRGTTLTKLTEDRQIR 103
QY 61 QGLEBELQKVLPGGDTYMHGEPERASEQIYYENRGYRTASVITALTGELHEDLFFYSER 120
DB 104 QGLEBELQKVLPGGDTYMHGEPERASEQIYYENRGYRTASVITALTGELHEDLFFYSER 163
QY 121 EANRSRDLGAIIVYCVGVKDFNETQARIADSKDHVPVNDGFQALQGIHSIL 173
DB 164 EANRSRDLGAIIVYCVGVKDFNETQARIADSKDHVPVNDGFQALQGIHSIL 216
RESULT 8
ADI00554
ID ADI00554 standard; protein; 342 AA.
XX
AC ADI00554;
XX
DT 22-APR-2004 (first entry)
XX
DE Human TANGO 197 FLAG fusion protein - plasmid p0613.
XX
KM fusion; von Willebrand factor A-like domain; vWF; antibacterial;
KM cutaneous; inhalation anthrax; human; TANGO 197 FLAG fusion; mutant;
KM plasmid p0613; mutain.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN US2003144193-A1.
XX
PD 31-JUL-2003.
XX
PF 24-JUL-2002; 2002US-00201292.
XX
PR 20-DEC-2001; 2001US-00038307.
XX
PA (ROTT/) ROTTMAN J B.
PA (OKEE/) O'KEEFE T L.
PA (OZKA/) OZKAYNAK E.
PA (HEAL/) HEALEY J J.
XX
PI Rottman JB, O'Keefe TL, Ozkaynak E, Healey JJ;
XX
DR WPI; 2003-720708/68.
XX
DR N-PSDB; ADI00553.
XX
PT New TANGO 197 and/or TANGO 216 fusion polypeptides for preventing or
PT ameliorating symptoms of anthrax comprises a von Willebrand factor A-like
PT domain (vWF) amino acid sequence and an amino acid sequence heterologous

PT to the vWF.
XX
XX Claim 45; SEQ ID NO 22; 86pp; English.
XX
XX
CC The invention relates to a novel fusion polypeptide comprising a von
CC Willebrand factor A-like domain (vWF) amino acid sequence and an amino
CC acid sequence heterologous to the vWF. The polypeptide of the invention
CC demonstrates antibacterial activities whilst the composition and method
CC may be useful in preventing or ameliorating the symptoms of cutaneous
CC and/or inhalation anthrax. The current sequence is that of the human
CC TANGO 197 FLAG fusion protein of the invention.
XX
SQ Sequence 342 AA;
SQ
Query Match 100.0%; Score 899; DB 7; Length 342;
Best Local Similarity 100.0%; Pred. No. 2.9e-92;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DLVPIIDKSGSVLHNMNIYFVQOLAHKFTSPQLRMSFTVFSRGTTLTKLTEDRQIR 60
DB 53 DLVPIIDKSGSVLHNMNIYFVQOLAHKFTSPQLRMSFTVFSRGTTLTKLTEDRQIR 112
QY 61 QGLEBELQKVLPGGDTYMHGEPERASEQIYYENRGYRTASVITALTGELHEDLFFYSER 120
DB 113 QGLEBELQKVLPGGDTYMHGEPERASEQIYYENRGYRTASVITALTGELHEDLFFYSER 172
QY 121 EANRSRDLGAIIVYCVGVKDFNETQARIADSKDHVPVNDGFQALQGIHSIL 173
DB 173 EANRSRDLGAIIVYCVGVKDFNETQARIADSKDHVPVNDGFQALQGIHSIL 225
RESULT 9
ADM64580
ID ADM64580 standard; protein; 342 AA.
XX
AC ADM64580;
XX
DT 03-JUN-2004 (first entry)
XX
DE Human TANGO197-FLAG epitope fusion protein.
XX
KM antibacterial; gene therapy;
KM von Willebrand factor A-like domain amino acid sequence;
KM vWF amino acid sequence; anthrax; *Bacillus anthracis*; cutaneous anthrax;
KM inhalation anthrax; human; TANGO197; FLAG; fusion protein; gene.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN US2003134786-A1.
XX
PD 17-JUL-2003.
XX
PF 20-DEC-2001; 2001US-00038307.
XX
PR 20-DEC-2001; 2001US-00038307.
XX
PA (ROTT/) ROTTMAN J B.
PA (OKEE/) O'KEEFE T L.
PA (OZKA/) OZKAYNAK E.
PA (HEAL/) HEALEY J J.
XX
PI Rottman JB, O'Keefe TL, Ozkaynak E, Healey JJ;
XX
DR WPI; 2003-829643/77.
XX
DR N-PSDB; ADM64579.
XX
PT New TANGO 197 and/or TANGO 216 fusion polypeptides for preventing or
PT ameliorating symptoms of anthrax comprises a von Willebrand factor A-like
PT domain (vWF) amino acid sequence and an amino acid sequence heterologous
PT to the vWF.
XX
XX Claim 44; SEQ ID NO 22; 64pp; English.

XX The invention describes a fusion polypeptide comprising a von Willebrand
CC factor A-like domain (vWF) amino acid sequence and an amino acid sequence
CC heterologous to the vWF. Also described are: a method of preventing or
CC ameliorating a symptom of anthrax in a subject thought to be at risk for
CC exposure to or suspected of having been exposed to Bacillus anthracis;
CC and a pharmaceutical composition comprising the novel fusion polypeptide.
CC The composition and method are useful in preventing or ameliorating
CC symptoms of cutaneous and/or inhalation anthrax. This is the amino acid
CC sequence of a fusion protein comprising human TANGO197 and FLAG that can
CC be used to treat exposure to or prevent a symptom of anthrax.

XX
SQ Sequence 342 AA;

Query Match 100.0%; Score 899; DB 7; Length 342;

Best Local Similarity 100.0%; Pred. No. 2,9e-92;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLVFIIDKSGSVLHHNNEIYFVEQLAHKFIISPOLRMSFIVFSTRGTTLMKLTEDREQIR 60
DB 53 DLVFIIDKSGSVLHHNNEIYFVEQLAHKFIISPOLRMSFIVFSTRGTTLMKLTEDREQIR 112
QY 61 QGLEBLQKVLPGGDTYMHGFERASEQIYYENRQGYRTASVITALTDEGLHEDLFFYSER 120
DB 113 QGLEBLQKVLPGGDTYMHGFERASEQIYYENRQGYRTASVITALTDEGLHEDLFFYSER 172
QY 121 EANRSRDLGAIIVYCVGVKDFNETOLARIADSKDHVPVNDGFQALQGIHSIL 173
DB 173 EANRSRDLGAIIVYCVGVKDFNETOLARIADSKDHVPVNDGFQALQGIHSIL 225

RESULT 10

AD100556 standard; protein; 345 AA.

XX
AC AD100556;

DT 22-APR-2004 (first entry)

DE Human TANGO 197 Histag fusion protein - plasmid p0614.

XX fusion; von Willebrand factor A-like domain; vWF; antibacterial;
KM cutaneous; inhalation anthrax; human; TANGO 197 Histag fusion; mutant;
KM plasmid p0614; mutain.

XX Homo sapiens.
OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 331 /note= "Wild-type Cys replaced by Ser followed by
FT thrombin cleavage site and His tag"

PN US2003144193-A1.

PD 31-JUL-2003.

PF 24-JUL-2002; 2002US-00201292.

PR 20-DEC-2001; 2001US-00038307.

XX (ROTT/) ROTTMAN J B.
PA (OKEE/) O'KEEFE T L.
PA (OZKA/) OZKAYNAK E.
PA (HEAL/) HEALEY J J.

PI Rottman JB, O'Keefe TL, Ozkaynak E, Healey JJ;

DR WPI; 2003-720708/68.

DR N-PSDB; AD100555.

XX New TANGO 197 and/or TANGO 216 fusion polypeptides for preventing or
PT ameliorating symptoms of anthrax comprises a von Willebrand factor A-like

PT domain (vWF) amino acid sequence and an amino acid sequence heterologous
PT to the vWF.

XX Claim 45; SEQ ID NO 24; 86pp; English.

XX The invention relates to a novel fusion polypeptide comprising a von
CC Willebrand factor A-like domain (vWF) amino acid sequence and an amino
CC acid sequence heterologous to the vWF. The polypeptide of the invention
CC demonstrates antibacterial activities whilst the composition and method
CC may be useful in preventing or ameliorating the symptoms of cutaneous
CC and/or inhalation anthrax. The current sequence is that of the human
CC TANGO 197 Histag fusion protein of the invention.

XX
SQ Sequence 345 AA;

Query Match 100.0%; Score 899; DB 7; Length 345;

Best Local Similarity 100.0%; Pred. No. 2,9e-92;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLVFIIDKSGSVLHHNNEIYFVEQLAHKFIISPOLRMSFIVFSTRGTTLMKLTEDREQIR 60
DB 44 DLVFIIDKSGSVLHHNNEIYFVEQLAHKFIISPOLRMSFIVFSTRGTTLMKLTEDREQIR 103
QY 61 QGLEBLQKVLPGGDTYMHGFERASEQIYYENRQGYRTASVITALTDEGLHEDLFFYSER 120
DB 104 QGLEBLQKVLPGGDTYMHGFERASEQIYYENRQGYRTASVITALTDEGLHEDLFFYSER 163
QY 121 EANRSRDLGAIIVYCVGVKDFNETOLARIADSKDHVPVNDGFQALQGIHSIL 173
DB 164 EANRSRDLGAIIVYCVGVKDFNETOLARIADSKDHVPVNDGFQALQGIHSIL 216

RESULT 11

ADM64582 standard; protein; 345 AA.

XX
AC ADM64582;

DT 03-JUN-2004 (first entry)

DE Human TANGO197-His tag fusion protein #1.

XX antibacterial; gene therapy;
KM von Willebrand factor A-like domain amino acid sequence;
KM vWF amino acid sequence; anthrax; Bacillus anthracis; cutaneous anthrax;
KM inhalation anthrax; human; TANGO197; his tag; fusion protein; gene.

XX Homo sapiens.
OS Synthetic.

PN US2003134786-A1.

PD 17-JUL-2003.

PF 20-DEC-2001; 2001US-00038307.

PR 20-DEC-2001; 2001US-00038307.

XX (ROTT/) ROTTMAN J B.
PA (OKEE/) O'KEEFE T L.
PA (OZKA/) OZKAYNAK E.
PA (HEAL/) HEALEY J J.

PI Rottman JB, O'Keefe TL, Ozkaynak E, Healey JJ;

DR WPI; 2003-829643/77.

DR N-PSDB; ADM64581.

XX New TANGO 197 and/or TANGO 216 fusion polypeptides for preventing or
PT ameliorating symptoms of anthrax comprises a von Willebrand factor A-like
PT domain (vWF) amino acid sequence and an amino acid sequence heterologous
PT to the vWF.

PS Claim 44, SEQ ID NO 24, 64pp; English.

XX The invention describes a fusion polypeptide comprising a von Willebrand
CC factor A-like domain (vWF) amino acid sequence and an amino acid sequence
CC heterologous to the vWF. Also described are: a method of preventing or
CC ameliorating a symptom of anthrax in a subject thought to be at risk for
CC exposure to or suspected of having been exposed to *Bacillus anthracis*;
CC and a pharmaceutical composition comprising the novel fusion polypeptide.
CC The composition and method are useful in preventing or ameliorating
CC symptoms of cutaneous and/or inhalation anthrax. This is the amino acid
CC sequence of a fusion protein comprising human TANGO197, thrombin cleavage
CC site and his tag that can be used to treat exposure to or prevent a
CC symptom of anthrax.

XX
XX Sequence 345 AA:

Query Match 100.0%; Score 899; DB 7; Length 345;
Best Local Similarity 100.0%; Pred. No. 2,9e-92;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLYFIIDKSGSVLHNNIYFVYEQLAHKFISPOLRMSFIVFSTRTTLTKLTEDRQIR 60
DB 44 DLYFIIDKSGSVLHNNIYFVYEQLAHKFISPOLRMSFIVFSTRTTLTKLTEDRQIR 103
QY 61 OGLEELQKVLPGDDTYHGEFPERASQIYYENRQGYRTASVITALTDEGLHEDLFFYSER 120
DB 104 OGLEELQKVLPGDDTYHGEFPERASQIYYENRQGYRTASVITALTDEGLHEDLFFYSER 163
QY 121 EANRSRDLAGIYVCVGVKDFENETOLARIADSKDHFVFNNGFOLQGIHSIL 173
DB 164 EANRSRDLAGIYVCVGVKDFENETOLARIADSKDHFVFNNGFOLQGIHSIL 216

RESULT 12
ID ABP54903 standard; protein: 368 AA.
XX
XX ABP54903;
AC
XX
DT 08-JAN-2003 (first entry)
XX
DB Human anthrax toxin receptor.
XX
XX Anthrax; toxin; receptor; human; antibacterial.
XX
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT 1..27
FT Peptide /label= signal_peptide
FT Region 27..321
FT /note= "PA-binding fragment, begins at any amino acid in
FT the range 27-43 and ends at any amino acid in the range
FT 221-321, region specifically described in Claim 3"
FT Protein 28..368
FT /label= Mature_protein
FT 28..320
FT /note= "extracellular domain"
FT 44..216
FT /note= "von Willebrand factor A domain"
FT 50
FT /note= "forms metal ion-dependent adhesion site (MIDAS)
FT motif with amino acid residues 52, 54, 118 and 150"
FT 52
FT /note= "forms metal ion-dependent adhesion site (MIDAS)
FT motif with amino acid residues 50, 54, 118 and 150"
FT 54
FT /note= "forms metal ion-dependent adhesion site (MIDAS)
FT motif with amino acid residues 50, 52, 118 and 150"
FT 118
FT /note= "forms metal ion-dependent adhesion site (MIDAS)
FT motif with amino acid residues 50, 52, 54 and 150"
FT 150
FT Region

FT /note= "forms metal ion-dependent adhesion site (MIDAS)
FT motif with amino acid residues 50, 52, 54 and 118"
FT 320..343
FT /note= "putative transmembrane domain"
FT Domain 344..368
FT /note= "cytoplasmic domain"
PN WO200246228-A2.
XX
XX 13-JUN-2002.
PD
XX
XX 03-OCT-2001; 2001MO-US030941.
PF
XX
XX 05-DEC-2000; 2000US-0251481P.
PR
XX
XX (MISC) WISCONSIN ALUMNI RES FOUND.
PA
PI Young JAT, Bradley KA, Collier RJ, Mogridge JS,
XX
XX WPI; 2002-713235/77.
DR
DR N-PSDB; ABV73881.
XX
XX Novel isolated polypeptide useful for identifying agent that prevents or
PT reduces effect of anthrax toxin on host cell, for treating human or non-
PT human animal suffering from anthrax.
XX
XX Claim 1; Page 29-30; 45pp; English.

CC The present sequence is the protein sequence of a human surface- bound
CC anthrax toxin receptor (ATR), as predicted from an isolated cDNA clone.
CC Anthrax toxin protective antigen (PA) binds to the ATR at a von
CC Willebrand factor A domain located in the extracellular domain of ATR.
CC The invention provides ATR polypeptides and polynucleotides, vectors,
CC host cells, and transgenic and knock-out animals. It also provides
CC methods for identifying molecules that bind the ATR and which reduce the
CC toxicity of anthrax toxin. A claimed method for treating anthrax in a
CC human or animal involves administering an agent that inhibits binding
CC between PA and ATR at a level effective to reduce the severity of
CC anthrax. Suitable agents include ATR or a PA-binding fragment of ATR, a
CC PA-binding polypeptide at least 80% identical to these, a fusion protein,
CC a monoclonal or polyclonal antibody, a polysaccharide, a lipid or a
CC nucleic acid
XX
XX
SQ Sequence 368 AA;

Query Match 100.0%; Score 899; DB 5; Length 368;
Best Local Similarity 100.0%; Pred. No. 3,2e-92;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLYFIIDKSGSVLHNNIYFVYEQLAHKFISPOLRMSFIVFSTRTTLTKLTEDRQIR 60
DB 44 DLYFIIDKSGSVLHNNIYFVYEQLAHKFISPOLRMSFIVFSTRTTLTKLTEDRQIR 103
QY 61 OGLEELQKVLPGDDTYHGEFPERASQIYYENRQGYRTASVITALTDEGLHEDLFFYSER 120
DB 104 OGLEELQKVLPGDDTYHGEFPERASQIYYENRQGYRTASVITALTDEGLHEDLFFYSER 163
QY 121 EANRSRDLAGIYVCVGVKDFENETOLARIADSKDHFVFNNGFOLQGIHSIL 173
DB 164 EANRSRDLAGIYVCVGVKDFENETOLARIADSKDHFVFNNGFOLQGIHSIL 216

RESULT 13
ID ADM64586 standard; protein: 384 AA.
XX
XX ADM64586;
AC
XX
DT 03-JUN-2004 (first entry)
XX
DE TANGO197 extracellular domain-mutant IGG Fc fusion protein #1.
XX
XX antibacterial; gene therapy;

KM von Willebrand factor A-like domain amino acid sequence;
KW vWF amino acid sequence; anthrax; Bacillus anthracis; cutaneous anthrax;
KM Inhalation anthrax; human; TANGO197; immunoglobulin G, IgG;
KM fragment of crystallisation; Fc.
XX
OS Homo sapiens.
OS Synthetic.
XX
XX US2003134786-A1.
XX
XX 17-JUL-2003.
XX
XX 20-DEC-2001; 2001US-00038307.
XX
XX 20-DEC-2001; 2001US-00038307.
XX
XX (ROT/) ROTTMAN J B.
PA (OXEE/) O'KEEFE T L.
PA (OZKA/) OZKAYNAK E.
PA (HEAL/) HEALEY J J.
XX
XX Rotman JB, O'Keefe TL, Ozkaynak E, Healey JJ;
XX
XX WPI; 2003-829643/77.
DR N-PsDB; ADM64585.
XX
XX New TANGO 197 and/or TANGO 216 fusion polypeptides for preventing or
PT ameliorating symptoms of anthrax comprises a von Willebrand factor A-like
PT domain (vWF) amino acid sequence and an amino acid sequence heterologous
PT to the vWF.
XX
XX
XX Claim 44, SEQ ID NO 10; 64pp; English.
XX
XX The invention describes a fusion polypeptide comprising a von Willebrand
CC factor A-like domain (vWF) amino acid sequence and an amino acid sequence
CC heterologous to the vWF. Also described are: a method of preventing or
CC ameliorating a symptom of anthrax in a subject thought to be at risk for
CC exposure to or suspected of having been exposed to Bacillus anthracis;
CC and a pharmaceutical composition comprising the novel fusion polypeptide.
CC The composition and method are useful in preventing or ameliorating
CC symptoms of cutaneous and/or inhalation anthrax. This is the amino acid
CC sequence of a fusion protein comprising human TANGO197 extracellular
CC region, minus the DG residues closest to the transmembrane region, and
CC immunoglobulin G (IgG) fragment of crystallisation (Fc) with mutations
CC L235A and G237A.
CC
XX
XX Sequence 384 AA:
SQ
Query Match 100.0%; Score 899; DB 7; Length 384;
Best Local Similarity 100.0%; Pred. No. 3,4e-92;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DLYFIIDKSGSVLHHNNEIYYPVQLAKRISPOLRMSFIVFSTRGTTLMKLTEDREQR 60
DB 44 DLYFIIDKSGSVLHHNNEIYYPVQLAKRISPOLRMSFIVFSTRGTTLMKLTEDREQR 103
QY 61 QGLEELQKVLPGSDTYNHEGFEPASQEIYYENRQGYRTASVITALTDEGLHEDLPFYSER 120
DB 104 QGLEELQKVLPGSDTYNHEGFEPASQEIYYENRQGYRTASVITALTDEGLHEDLPFYSER 163
QY 121 EARRSDILGAIYVCVQKDFENFOLARIASKHVPVNGPQALOGIHSIL 173
DB 164 EARRSDILGAIYVCVQKDFENFOLARIASKHVPVNGPQALOGIHSIL 216
RESULT 14
AAE01439
ID AAE01439 standard; protein; 403 AA.
XX
XX AAE01439;
XX
XX 17-JUL-2001 (first entry)
XX

DE Human gene 4 encoded secreted protein HMLFR02, SEQ ID NO:94.
XX
XX Human; secreted protein; proliferative disorder; cancer; tumour;
XX foetal abnormality; developmental abnormality; haemopoietic disorder;
XX immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
XX inflammation; allergy; neurological disorder; Alzheimer's disease;
XX Parkinson's disease; cognitive disorder; schizophrenia; asthma;
XX skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
XX cardiovascular disorder; angiogenic disorder; kidney disorder;
XX gastrointestinal disorder; pregnancy-related disorder;
XX endocrine disorder; infection; wound healing; vulvovaginal; cell culture;
XX chemotaxis; food additive; gene therapy; binding partner identification;
XX chromosome 19.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
FH 1..27
FT Peptide /label=Signal_peptide
FT 28..403
FT Protein /note="Mature human secreted protein"
XX
XX MO200134626-A1.
XX
XX 17-MAY-2001.
XX
XX 01-NOV-2000; 2000MO-US030045.
XX
XX 05-NOV-1999; 99US-0163581P.
XX 30-JUN-2000; 2000US-0215133P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Komatsoulis GA, Moore PA, Birse CE, Ni J;
XX WPI; 2001-308778/32.
XX N-PsDB; AAD05303.
XX
XX New nucleic acid molecules encoding 28 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives.
XX
XX Claim 11; Page 485-486; 562pp; English.
XX
XX AAD05300-AAD05379 represent cDNAs corresponding to 28 human secreted
CC protein genes, and AAE01436-AAE01513 represent the proteins they encode.
CC AAB01514-AAE01544 represent human secreted protein fragments or variants.
CC The genes and their secreted proteins are useful for preventing, treating
CC or ameliorating medical conditions, e.g., by protein or gene therapy.
CC Pathological conditions can be diagnosed by determining the amount of the
CC new protein in a sample or by determining the presence of mutations in
CC the new genes. Specific uses are described for each of the 28 genes,
CC based on the tissues in which they are most highly expressed, and include
CC developing products for the diagnosis and treatment of proliferative
CC disorders, cancer, tumours, foetal and developmental abnormalities,
CC haemopoietic disorders, diseases of the immune system, AIDS, autoimmune
CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,
CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
CC angiogenic disorders, kidney disorders, gastrointestinal disorders,
CC pregnancy-related disorders, endocrine disorders, and infections. The
CC proteins can also be used to aid wound healing and epithelial cell
CC proliferation, to prevent skin aging due to sunburn, to maintain organs
CC before transplantation, for supporting cell culture of primary tissues,
CC to regenerate tissues, to identify their cognate ligands or binding
CC partners, and in chemotaxis, and can be used as a food additive or
CC preservative to modify storage properties. Antibodies specific for a
CC protein of the invention can be used in alleviating symptoms associated
CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,
CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The
CC present sequence represents a human secreted protein of the invention
XX

80 Sequence 403 AA;
Query Match 100.0%; Score 899; DB 4; Length 403;
Best Local Similarity 100.0%; Pred. No. 3.7e-92;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DLVFIIDKSGSVLHNNHNIYFVEQLAHKFTSPQLRMSFIVFSTRTTLMKLTEDRQIR 60
DB 44 DLVFIIDKSGSVLHNNHNIYFVEQLAHKFTSPQLRMSFIVFSTRTTLMKLTEDRQIR 103
QY 61 QGEBELQKVLPGDPTVNHGFERASEQIYYENRGYRTASVIALTDGELHEDLFFYSER 120
DB 104 QGEBELQKVLPGDPTVNHGFERASEQIYYENRGYRTASVIALTDGELHEDLFFYSER 163
QY 121 EANRSRLGAIIVYCVGVKDFNETQLARIADSKDHVPVNDGFQALQGIHISIL 173
DB 164 EANRSRLGAIIVYCVGVKDFNETQLARIADSKDHVPVNDGFQALQGIHISIL 216
RESULT 15
ABG63874
ID ABG63874 standard; protein; 403 AA.
XX
AC ABG63874;
XX
DT 27-AUG-2002 (first entry)
XX
DE Human albumin fusion protein #549.
XX
KW Albumin fusion protein; therapeutic protein X; human albumin; HA;
KW human serum albumin; HSA; cancer; reproductive disorder;
KW digestive disorder; immune disorder; endocrine disorder;
KW hematopoietic disorder; neural disorder; connective disorder;
KW cytoskeletal; antifertility; antiinflammatory; antiulcer;
KW immunomodulator; anti-HIV; antidiabetic; haemostatic; noctropic;
KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
KW osteopathic; antichratic.
XX
OS Homo sapiens.
XX
SS Synthetic.
XX
PN MO200177137-A1.
XX
PD 18-OCT-2001.
XX
PF 12-APR-2001; 2001WO-US011988.
XX
PR 12-APR-2000; 2000US-0229358P.
XX
PR 25-APR-2000; 2000US-0199384P.
XX
PR 21-DEC-2000; 2000US-0256931P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Haseltine WA;
XX
DR WPI; 2002-010886/01.
XX
PT New fusion protein for treating disease e.g. diabetes comprises an
XX albumin fused to a therapeutic protein.
XX
PS Claim 1; Page 874-875; 2102pp; English.
XX
CC The present invention relates to albumin fusion proteins comprising a
CC therapeutic protein X and human albumin (HA, also known as human serum
CC albumin, HSA). The proteins are useful for treating a disease or disorder
CC that may be modulated by therapeutic protein X. The albumin extends the
CC shelf-life of protein X, and may increase its biological in vitro/in vivo
CC activity. The protein is useful for treating and diagnosing disorders
CC such as cancer, reproductive disorders, digestive disorders (e.g. Crohn's
CC disease, ulcerative colitis), immune disorders (e.g. acquired
CC immunodeficiency syndrome, AIDS), endocrine disorders (e.g. diabetes),
CC hematopoietic disorders, neural disorders (e.g. Alzheimer's,
CC Parkinson's, Creutzfeldt-Jacob disease, encephalomyelitis, meningitis,

CC schizophrenia), and connective disorders (e.g. osteoporosis, arthritis).
CC ABG63326-ABG65518 represent albumin fusion proteins of the invention
XX
80 Sequence 403 AA;
Query Match 100.0%; Score 899; DB 5; Length 403;
Best Local Similarity 100.0%; Pred. No. 3.7e-92;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DLVFIIDKSGSVLHNNHNIYFVEQLAHKFTSPQLRMSFIVFSTRTTLMKLTEDRQIR 60
DB 44 DLVFIIDKSGSVLHNNHNIYFVEQLAHKFTSPQLRMSFIVFSTRTTLMKLTEDRQIR 103
QY 61 QGEBELQKVLPGDPTVNHGFERASEQIYYENRGYRTASVIALTDGELHEDLFFYSER 120
DB 104 QGEBELQKVLPGDPTVNHGFERASEQIYYENRGYRTASVIALTDGELHEDLFFYSER 163
QY 121 EANRSRLGAIIVYCVGVKDFNETQLARIADSKDHVPVNDGFQALQGIHISIL 173
DB 164 EANRSRLGAIIVYCVGVKDFNETQLARIADSKDHVPVNDGFQALQGIHISIL 216

Search completed: June 13, 2005, 19:56:53
Job time : 65.0514 secs

THIS PAGE BLANK (USPTO)

Db 142 DIAFLIDSGSINORDPAQMKDFVKALMGEPASTSLFSLMQYSNLKHTFTPEKNIL 201
Qy 55 DREQIRQGLEBELQKVLPGSDTYMHGEPERASEQIYYENRQGYRTA-SVIALTDGELHED 113
Db 202 DQSLVDPIVQLQ-----GLTYTATGIRTYMBELFHSKNGSRKSAKKILLVITDGQKYRD 256
Qy 114 LFFYSE--REANRSDIGAIVYCVGYKD-FNE-TQLARI-----ADSKDHVPYNDGFOA 164
Db 257 PLEYSDVTPAADKA---GIRYALGVGDAPFOEFTALKELNTIGSAPPQDHVFKVGN-FAA 312
Qy 165 LQGI 168
Db 313 LRSI 316

RESULT 2
US-08-485-618-37
; Sequence 37, Application US/08485618
; Patent No. 5728533
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vlieten, Monica
; TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,618
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32797
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1151 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-485-618-37

Query Match 15.4%; Score 138.5; DB 1; Length 1151;
Best Local Similarity 28.3%; Pred. No. 6.6e-07;
Matches 52; Conservative 36; Mismatches 71; Indels 25; Gaps 10;
Qy 1 DLVFIIDKSGSV-LHNMNIYYFVEQIAHKFISPOLRMSPIVSTGTTLMLKLT-----54
Db 142 DIAFLIDSGSINORDPAQMKDFVKALMGEPASTSLFSLMQYSNLKHTFTPEKNIL 201

Qy 55 DREQIRQGLEBELQKVLPGSDTYMHGEPERASEQIYYENRQGYRTA-SVIALTDGELHED 113
Db 202 DQSLVDPIVQLQ-----GLTYTATGIRTYMBELFHSKNGSRKSAKKILLVITDGQKYRD 256
Qy 114 LFFYSE--REANRSDIGAIVYCVGYKD-FNE-TQLARI-----ADSKDHVPYNDGFOA 164
Db 257 PLEYSDVTPAADKA---GIRYALGVGDAPFOEFTALKELNTIGSAPPQDHVFKVGN-FAA 312
Qy 165 LQGI 168
Db 313 LRSI 316

RESULT 3
US-08-362-652-37
; Sequence 37, Application US/08362652
; Patent No. 5766850
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vlieten, Monica
; TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,652
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32391
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1151 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-362-652-37

Query Match 15.4%; Score 138.5; DB 1; Length 1151;
Best Local Similarity 28.3%; Pred. No. 6.6e-07;
Matches 52; Conservative 36; Mismatches 71; Indels 25; Gaps 10;
Qy 1 DLVFIIDKSGSV-LHNMNIYYFVEQIAHKFISPOLRMSPIVSTGTTLMLKLT-----54
Db 142 DIAFLIDSGSINORDPAQMKDFVKALMGEPASTSLFSLMQYSNLKHTFTPEKNIL 201
Qy 55 DREQIRQGLEBELQKVLPGSDTYMHGEPERASEQIYYENRQGYRTA-SVIALTDGELHED 113
Db 202 DQSLVDPIVQLQ-----GLTYTATGIRTYMBELFHSKNGSRKSAKKILLVITDGQKYRD 256
Qy 114 LFFYSE--REANRSDIGAIVYCVGYKD-FNE-TQLARI-----ADSKDHVPYNDGFOA 164

Db 257 PLEYSVDIIPADKA--GIIRYAI GVDARFOEPALKEMLNTIGSAPPODHVFKVGN-FAA 312
QY 165 LOGI 168
Db 313 LRSI 316

RESULT 4

US-08-605-672-37
Sequence 37, Application US/08605672
Patent No. 5817515
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vlieten, Monica
TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Seer Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/605,672
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 1151 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-605-672-37

Query Match 15.4%; Score 138.5; DB 2; Length 1151;
Best Local Similarity 28.3%; Pred. No. 6.6e-07;
Matches 52; Conservative 36; Mismatches 71; Indels 25; Gaps 10;
QY 1 DLYFIIDKSGSV-LHMHNEIYYFVEQLAHKRFISPOLMSFIVSTRTGTLMLKLE----- 54
Db 142 DIAFLIDGSSISINORDPAQKMDFFKALMGEPASTSTLSLMQYSNLIKHTFTFEPKNIL 201
QY 55 DREQIRGLBELQKVLPGSGDTYMHGEPERASEQIYYENRQGYRTA-SVIALTLTGELHED 113
Db 202 DPOSIVDPIVOLQ-----GLTYATGIRTYWELFHSKNGSRKSAKKILLVITDGQRYRD 256
QY 114 LRFYSE--REANSRDGAIVCYGVKD-FNE-TOLARI-----ADSKDHVPVNDGFOA 164
Db 257 PLEYSVDIIPADKA--GIIRYAI GVDARFOEPALKEMLNTIGSAPPODHVFKVGN-FAA 312

QY 165 LOGI 168
Db 313 LRSI 316

RESULT 5

US-08-482-293A-37
Sequence 37, Application US/08482293A
Patent No. 5831029
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vlieten, Monica
TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Seer Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,293A
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 1151 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-482-293A-37

Query Match 15.4%; Score 138.5; DB 2; Length 1151;
Best Local Similarity 28.3%; Pred. No. 6.6e-07;
Matches 52; Conservative 36; Mismatches 71; Indels 25; Gaps 10;
QY 1 DLYFIIDKSGSV-LHMHNEIYYFVEQLAHKRFISPOLMSFIVSTRTGTLMLKLE----- 54
Db 142 DIAFLIDGSSISINORDPAQKMDFFKALMGEPASTSTLSLMQYSNLIKHTFTFEPKNIL 201
QY 55 DREQIRGLBELQKVLPGSGDTYMHGEPERASEQIYYENRQGYRTA-SVIALTLTGELHED 113
Db 202 DPOSIVDPIVOLQ-----GLTYATGIRTYWELFHSKNGSRKSAKKILLVITDGQRYRD 256
QY 114 LRFYSE--REANSRDGAIVCYGVKD-FNE-TOLARI-----ADSKDHVPVNDGFOA 164
Db 257 PLEYSVDIIPADKA--GIIRYAI GVDARFOEPALKEMLNTIGSAPPODHVFKVGN-FAA 312
QY 165 LOGI 168


```
/ NUMBER OF SEQ ID NOS: 114
/ SOFTWARE: Patentln Ver. 2.0
/ SEQ ID NO 37
/ LENGTH: 1151
/ TYPE: PRT
/ ORGANISM: Rattus rattus
/ NAME/KEY: misc_feature
/ LOCATION: 361
/ OTHER INFORMATION: Xaa = any or unknown amino acid
/ NAME/KEY: misc_feature
/ LOCATION: 464
/ OTHER INFORMATION: Xaa = any or unknown amino acid
/ NAME/KEY: misc_feature
/ LOCATION: 486
/ OTHER INFORMATION: Xaa = any or unknown amino acid
/ NAME/KEY: misc_feature
/ LOCATION: 506
/ OTHER INFORMATION: Xaa = any or unknown amino acid
/ NAME/KEY: misc_feature
/ LOCATION: 1117
/ OTHER INFORMATION: Xaa = any or unknown amino acid
/ NAME/KEY: misc_feature
/ LOCATION: 1118
/ OTHER INFORMATION: Xaa = any or unknown amino acid
/ US-09-688-307A-37
```

```
Query Match 15.4%; Score 138.5; DB 4; Length 1151;
Best Local Similarity 28.3%; Pred. NO. 6.6e-07;
Matches 52; Conservative 36; Mismatches 71; Indels 25; Gaps 10;
```

```
QY 1 DLVFIIDKSGSV-LHHNNEIYYFVEQLAHKFIISPOLMSFVSTRTTLMKLT----- 54
DB 142 DIAFLIDGSGSINGRDPQKMDPVKALMGEPASTSTFLSLQYSNLIKHTFTPEFNIL 201
QY 55 DREQIRGSELQKVLFGSDTYMHEGPERASEQIYYENRQGYRTA-SVITALTDELHED 113
DB 202 DQSLVDPIVOLQ-----GLTYATGIRTYMBELFHSKNGSRKSAKKILLVITDGQKXRD 256
QY 114 LFPYSE--REANSRDGAIVYCVGYD-FNE-TQLARI-----ADSKDHVPYNDGFOA 164
DB 257 PLEYSVDIIPADKA--GIIRYALVGDAFOEPALKELTIGSAPPQDHVKVGN-FAA 312
QY 165 LQGI 168
DB 313 LRGI 316
```

```
RESULT 9
US-09-350-259-37
/ Sequence 37, Application US/09350259
/ Patent No. 6620915
/ GENERAL INFORMATION:
/ APPLICANT: Gallatin, Michael W.
/ TITLE OF INVENTION: No. 6620915el Human 2
/ FILE REFERENCE: 27866/35004
/ CURRENT APPLICATION NUMBER: US/09/350,259
/ FILING DATE: 1999-07-08
/ EARLIER APPLICATION NUMBER: 09/193,043
/ EARLIER FILING DATE: 1998-11-16
/ EARLIER APPLICATION NUMBER: 08/173,497
/ EARLIER FILING DATE: 1993-12-23
/ EARLIER APPLICATION NUMBER: 08/286,889
/ EARLIER FILING DATE: 1994-08-05
/ EARLIER APPLICATION NUMBER: 08/362,652
/ EARLIER FILING DATE: 1994-12-21
/ EARLIER APPLICATION NUMBER: 08/943,363
/ EARLIER FILING DATE: 1997-10-03
/ NUMBER OF SEQ ID NOS: 114
/ SOFTWARE: Patentln Ver. 2.0
/ SEQ ID NO 37
/ LENGTH: 1151
```

```
/ TYPE: PRT
/ ORGANISM: Rattus rattus
/ US-09-350-259-37
```

```
Query Match 15.4%; Score 138.5; DB 4; Length 1151;
Best Local Similarity 28.3%; Pred. NO. 6.6e-07;
Matches 52; Conservative 36; Mismatches 71; Indels 25; Gaps 10;
```

```
QY 1 DLVFIIDKSGSV-LHHNNEIYYFVEQLAHKFIISPOLMSFVSTRTTLMKLT----- 54
DB 142 DIAFLIDGSGSINGRDPQKMDPVKALMGEPASTSTFLSLQYSNLIKHTFTPEFNIL 201
QY 55 DREQIRGSELQKVLFGSDTYMHEGPERASEQIYYENRQGYRTA-SVITALTDELHED 113
DB 202 DQSLVDPIVOLQ-----GLTYATGIRTYMBELFHSKNGSRKSAKKILLVITDGQKXRD 256
QY 114 LFPYSE--REANSRDGAIVYCVGYD-FNE-TQLARI-----ADSKDHVPYNDGFOA 164
DB 257 PLEYSVDIIPADKA--GIIRYALVGDAFOEPALKELTIGSAPPQDHVKVGN-FAA 312
QY 165 LQGI 168
DB 313 LRGI 316
```

```
RESULT 10
US-08-485-618-55
/ Sequence 55, Application US/08485618
/ Patent No. 5728533
/ GENERAL INFORMATION:
/ APPLICANT: Gallatin, W. Michael
/ TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit
/ NUMBER OF SEQUENCES: 103
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
/ STREET: 233 South Wacker Drive, 6300 Sear Tower
/ CITY: Chicago
/ STATE: Illinois
/ COUNTRY: United States
/ ZIP: 60606-6402
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentln Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/485,618
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/173,497
/ FILING DATE: 23-DEC-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/286,889
/ FILING DATE: 5-AUG-1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/362,652
/ FILING DATE: 21-DEC-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Williams Jr., Joseph A.
/ REGISTRATION NUMBER: 38,659
/ REFERENCE/DOCKET NUMBER: 27866/32797
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 312-474-6300
/ TELEFAX: 312-474-0448
/ TELEX: 25-3856
/ INFORMATION FOR SEQ ID NO: 55:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1161 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
```

US-08-485-618-55

Query Match	15.4%;	Score 138.5;	DB 1;	Length 1161;
Best Local Similarity	28.3%;	Pred. No. 6.7e-07;		
Matches 52;	Conservative 36;	Mismatches 71;	Indels 25;	Gaps 10;

[illegible]

RESULT 11

```

US-08-362-652-55
; Sequence 55, Application US/08362652
; Patent No. 5766850
;
GENERAL INFORMATION:
;
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vliet, Monica
TITLE OF INVENTION: No. 5766850e1 Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESSES:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borum
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/362,652
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/22391
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 1161 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-362-652-55

```

Query Match	15.4%	Score 138.5;	DB 1;	Length 1161;
Best Local Similarity	28.3%	Pred. No. 6.7e-07;		
Matches	52;	Conservative	36;	Mismatches 71;
			Indels	25;
			Gaps	10;

[illegible]

RESULT 12

```

US-08-605-672-55
: Sequence 55, Application US/08605672
: Patent No. 5817515
:
: GENERAL INFORMATION:
:
: APPLICANT: Gallatin, W. Michael
: APPLICANT: Van der Vlieten, Monica
: TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit
: NUMBER OF SEQUENCES: 103
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
: STREET: 233 South Wacker Drive, 6300 Sear Tower
: CITY: Chicago
: STATE: Illinois
: COUNTRY: United States
: ZIP: 60606-6402
:
: COMPUTER READABLE FORM:
:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/605,672
: FILING DATE:
:
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/173,497
: FILING DATE: 23-DEC-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/286,889
: FILING DATE: 5-AUG-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/362,652
: FILING DATE: 21-DEC-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Williams Jr., Joseph A.
: REGISTRATION NUMBER: 38,659
: REFERENCE/DOCKET NUMBER: 27866/32684
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 312-474-3100
: TELEFAX: 312-474-0448
:
: TELEX: 25-3856
: INFORMATION FOR SEQ ID NO: 55:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1161 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
:
: IS-08-605-672-55

```

Query Match	15.4%	Score 138.5;	DB 2;	Length 1161;
Best Local Similarity	28.34%	Pred. No. 6,7e-07;		
Matches	52;	Conservative	36;	Mismatches 71;
			Indels	25;
			Gaps	10;
1	DLYFIIDKSSV-LHHNNEIYYVEVEQLAHKRFISPOLMSFVSTRGCTLMKLTIE----	54		

Db 152 DIAPLIDSGSSINORDPAQKMDPVKALMGEPASSTLSLMQYNSILKTHFTPEFKNII 211
QY 55 DREQIROGLEBELQVLPDGTVMHEGFERASEQIYENROGYRTA-SVIALTDGELHED 113
Db 212 DPGLVDPIVQLQ-----GLTYTATGIRTWEEELFHSKNGSRKSAKILLVITDGOQRD 266
QY 114 LFFYSE--REANSRDLGALIVCYGVND-FNE-TQLARI-----ADSKDHFPVNDGFOA 164
Db 267 PLEYSDVIPADKX---GIIRYALGVGDAPQEPALKEMLTIGSAPQDHVKVGN-FAA 322
QY 165 LQGI 168
Db 323 LRST 326

RESULT 13
US-08-482-293A-55
Sequence 55, Application US/08482293A
Patent No. 5831029

GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael

APPLICANT: Van der Vlieten, Monica
TITLE OF INVENTION: No. 5831029e1 Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois

COUNTRY: United States
ZIP: 60606-6402

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,293A

FILING DATE:
CLASSIFICATION: 530

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497

FILING DATE: 23-DEC-1993
APPLICATION NUMBER: US 08/286,889

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889

FILING DATE: 5-AUG-1994
APPLICATION DATA:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652

FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:

NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659

REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448

TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 55:

SEQUENCE CHARACTERISTICS:
LENGTH: 1161 amino acids

TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-482-293A-55

Query Match 15.4%; Score 138.5; DB 2; Length 1161;
Best Local Similarity 28.3%; Pred. No. 6.7e-07;
Matches 52; Conservative 36; Mismatches 71; Indels 25; Gaps 10;

QY 1 DLVFIILKSGSV-LHHNNEIYFVEQLAHKFIISPOLMSFIVSRTGTTMLKLE----- 54
Db 152 DIAPLIDSGSSINORDPAQKMDPVKALMGEPASSTLSLMQYNSILKTHFTPEFKNII 211

QY 55 DREQIROGLEBELQVLPDGTVMHEGFERASEQIYENROGYRTA-SVIALTDGELHED 113
Db 212 DPGLVDPIVQLQ-----GLTYTATGIRTWEEELFHSKNGSRKSAKILLVITDGOQRD 266
QY 114 LFFYSE--REANSRDLGALIVCYGVND-FNE-TQLARI-----ADSKDHFPVNDGFOA 164
Db 267 PLEYSDVIPADKX---GIIRYALGVGDAPQEPALKEMLTIGSAPQDHVKVGN-FAA 322
QY 165 LQGI 168
Db 323 LRST 326

RESULT 14
US-08-943-363-55
Sequence 55, Application US/08943363
Patent No. 5837478

GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael

APPLICANT: Van der Vlieten, Monica
TITLE OF INVENTION: No. 5837478e1 Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois

COUNTRY: United States
ZIP: 60606-6402

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,363

FILING DATE:
CLASSIFICATION: 530

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497

FILING DATE: 23-DEC-1993
APPLICATION NUMBER: US 08/286,889

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889

FILING DATE: 5-AUG-1994
APPLICATION DATA:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652

FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:

NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659

REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448

TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 55:

SEQUENCE CHARACTERISTICS:
LENGTH: 1161 amino acids

TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-943-363-55

Query Match 15.4%; Score 138.5; DB 2; Length 1161;
Best Local Similarity 28.3%; Pred. No. 6.7e-07;
Matches 52; Conservative 36; Mismatches 71; Indels 25; Gaps 10;

QY 1 DLVFIILKSGSV-LHHNNEIYFVEQLAHKFIISPOLMSFIVSRTGTTMLKLE----- 54
Db 152 DIAPLIDSGSSINORDPAQKMDPVKALMGEPASSTLSLMQYNSILKTHFTPEFKNII 211
QY 55 DREQIROGLEBELQVLPDGTVMHEGFERASEQIYENROGYRTA-SVIALTDGELHED 113

```
Db      212 DPGSLVDPYQLQ-----GLTYTATGIRTWMBELFHSKNGSRKSAKKILLVITDQCKYRD 266
QY      114 LFFYSE--REANRSRDIGAIYVCVAKD-FNE-TQLARI-----ADSKDHVPVNDGFOA 164
Db      267 PLEYSVDYIPADKA---GIIRYAIGVGDAPFOEPTALKELNTIGSAPPODHVFKVGN-FAA 322
QY      165 LOGI 168
Db      323 LRSI 326
```

RESULT 15

```
US-09-193-043-55
; Sequence 55: Application US/09193043
; Patent No. 6251395
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6251395el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/193, 043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173, 497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286, 889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362, 652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943, 363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 55
; LENGTH: 1161
; TYPE: PRT
; ORGANISM: Rattus rattus
US-09-193-043-55
```

```
Query Match      15.4%; Score 138.5; DB 3; Length 1161;
Best Local Similarity 28.3%; Pred. No. 6,7e-07;
Matches 52; Conservative 36; Mismatches 71; Indels 25; Gaps 10;
```

```
QY      1 DLYFIIDKSGSV-LHMHNEIYFVEQLAHKEISPOLMSFIVPSTRGITLMKLT----- 54
Db      152 DIAFLIDGSGSINORDPAQMKDFFKALMGEPASTSTLPSLMQYSNLIKTHFTFEKNIL 211
QY      55 DREQIRQGLELQKVLPGDVTYMEHGFERASEQIYENROGYRTA-SVIALTDGELHED 113
Db      212 DPGSLVDPYQLQ-----GLTYTATGIRTWMBELFHSKNGSRKSAKKILLVITDQCKYRD 266
QY      114 LFFYSE--REANRSRDIGAIYVCVAKD-FNE-TQLARI-----ADSKDHVPVNDGFOA 164
Db      267 PLEYSVDYIPADKA---GIIRYAIGVGDAPFOEPTALKELNTIGSAPPODHVFKVGN-FAA 322
QY      165 LOGI 168
Db      323 LRSI 326
```

Search completed: June 13, 2005, 20:05:09
Job time : 16.9024 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 20:03:29 ; Search time 55.4703 Seconds
(without alignments)
1195.537 Million cell updates/sec

Title: US-09-970-076-2_COPY_44_216
Perfect score: 899
Sequence: 1 DLVFILDKSGSVLHNMNEIY.....HFPVNDGFQALQGIHSIL 173

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1710399 seqs, 383334425 residues

Total number of hits satisfying chosen parameters: 1710399

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubppaa/PCR_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubppaa/PCRUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubppaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubppaa/US10B_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubppaa/US10B_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubppaa/US10B_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubppaa/US11_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/1/pubppaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	899	100.0	328	US-10-038-307-26	Sequence 26, App1
2	899	100.0	328	US-10-201-292-26	Sequence 26, App1
3	899	100.0	333	US-09-796-753-12	Sequence 12, App1
4	899	100.0	333	US-10-038-307-2	Sequence 2, App1
5	899	100.0	333	US-10-201-292-2	Sequence 22, App1
6	899	100.0	342	US-10-038-307-22	Sequence 22, App1
7	899	100.0	342	US-10-201-292-22	Sequence 22, App1
8	899	100.0	345	US-10-038-307-24	Sequence 24, App1
9	899	100.0	345	US-10-201-292-24	Sequence 24, App1
10	899	100.0	403	US-09-833-245-621	Sequence 621, App1
11	899	100.0	460	US-10-201-292-28	Sequence 28, App1

12	899	100.0	460	US-10-201-292-30	Sequence 30, App1
13	899	100.0	479	US-10-201-292-32	Sequence 32, App1
14	899	100.0	504	US-10-201-292-34	Sequence 34, App1
15	899	100.0	529	US-10-201-292-36	Sequence 36, App1
16	899	100.0	538	US-10-047-542-99	Sequence 99, App1
17	899	100.0	551	US-10-038-307-18	Sequence 18, App1
18	899	100.0	551	US-10-201-292-18	Sequence 18, App1
19	899	100.0	564	US-09-918-715-187	Sequence 187, App
20	899	100.0	564	US-09-918-715-232	Sequence 232, App
21	899	100.0	564	US-10-038-307-20	Sequence 20, App1
22	899	100.0	564	US-10-201-292-20	Sequence 20, App1
23	899	100.0	564	US-10-301-292-199	Sequence 199, App
24	899	100.0	564	US-10-408-765A-1823	Sequence 1823, App
25	899	100.0	564	US-10-474-794-187	Sequence 187, App
26	899	100.0	564	US-10-474-794-232	Sequence 232, App
27	899	99.4	403	US-09-833-245-620	Sequence 620, App
28	899	99.0	562	US-09-918-715-194	Sequence 194, App
29	899	99.0	562	US-09-918-715-301	Sequence 301, App
30	899	99.0	562	US-10-474-794-194	Sequence 194, App
31	899	99.0	562	US-10-474-794-301	Sequence 301, App
32	878	97.7	534	US-10-038-307-12	Sequence 12, App1
33	878	97.7	534	US-10-201-292-12	Sequence 12, App1
34	878	97.7	543	US-10-038-307-10	Sequence 10, App1
35	878	97.7	543	US-10-038-307-14	Sequence 14, App1
36	878	97.7	543	US-10-038-307-16	Sequence 16, App1
37	878	97.7	543	US-10-201-292-10	Sequence 10, App1
38	878	97.7	543	US-10-201-292-14	Sequence 14, App1
39	878	97.7	543	US-10-201-292-16	Sequence 16, App1
40	519	57.7	488	US-10-038-307-6	Sequence 6, App1
41	519	57.7	488	US-10-201-292-6	Sequence 6, App1
42	519	57.7	488	US-10-368-087-16	Sequence 16, App1
43	519	57.7	488	US-10-104-047-2639	Sequence 2639, App
44	519	57.7	488	US-09-796-753-54	Sequence 54, App1
45	517	57.5	487	US-09-796-753-54	Sequence 54, App1

ALIGNMENTS

RESULT 1
US-10-038-307-26
Sequence 26, Application US/10038307
Publication No. US20030134786A1
GENERAL INFORMATION:
APPLICANT: James B. ROTTMAN
APPLICANT: Theresa L. O'KEEFE
APPLICANT: Engin OZKANNAK
TITLE OF INVENTION: Judith V. HEALEY
FILE REFERENCE: 7853-253-999
CURRENT APPLICATION NUMBER: US/10/038,307
CURRENT FILING DATE: 2002-06-28
NUMBER OF SEQ ID NOS: 26
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 26
LENGTH: 328
TYPE: PRT
ORGANISM: Homo sapiens
US-10-038-307-26

Query Match 100.0%; Score 899; DB 14; Length 328;
Best Local Similarity 100.0%; Pred. No. 1e-86;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLVFILDKSGSVLHNMNEIYFVEQLAKRISPOLRMSFIYFSGRTTLMKLTEDRQIR 60
DB 44 DLVFILDKSGSVLHNMNEIYFVEQLAKRISPOLRMSFIYFSGRTTLMKLTEDRQIR 103
QY 61 OGAEELQKVPKGGDTYHGEFPERASEQIYYENRGYVTASVITALTGELHEDFFYSER 120
DB 104 OGAEELQKVPKGGDTYHGEFPERASEQIYYENRGYVTASVITALTGELHEDFFYSER 163
QY 121 EANKSRDLGAIYVCVGVKDNRETOLARIADSKOHVFPVNDGFQALQGIHSIL 173

```
Db      164 EANRSDLGAIYCVGVKDFNETOLARIADSKDHVFPVNDGFQALOGIHSIL 216

RESULT 2
US-10-201-292-26
; Sequence 26, Application US/10201292
; Publication No. US2003014193A1
; GENERAL INFORMATION:
; APPLICANT: James B. ROTTMAN
; APPLICANT: Theresa L. O'KEEFE
; APPLICANT: Engin OZKAYNAK
; APPLICANT: Judith J. HEALEY
; TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
; FILE REFERENCE: 7853-253-999
; CURRENT APPLICATION NUMBER: US/10/201,292
; CURRENT FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-201-292-26

Query Match      100.0%; Score 899; DB 14; Length 328;
Best Local Similarity 100.0%; Pred. No. 1e-86;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 DLYFIIDKSGSVLHNNHNIYFYFVQLAHKFIISPOLRMSFVIFSTRGTTLMKLTEDRQIR 60
        |||||||
DB      44 DLYFIIDKSGSVLHNNHNIYFYFVQLAHKFIISPOLRMSFVIFSTRGTTLMKLTEDRQIR 103
        |||||||

QY      61 QGLEELQKVLPGSDTYMHGFEFASQIYYENRGYRTASVITALTDEGLHEDLFFYSER 120
        |||||||
DB      104 QGLEELQKVLPGSDTYMHGFEFASQIYYENRGYRTASVITALTDEGLHEDLFFYSER 163
        |||||||

QY      121 EANRSDLGAIYCVGVKDFNETOLARIADSKDHVFPVNDGFQALOGIHSIL 173
        |||||||
DB      164 EANRSDLGAIYCVGVKDFNETOLARIADSKDHVFPVNDGFQALOGIHSIL 216
        |||||||

RESULT 3
US-09-796-753-12
; Sequence 12, Application US/09796753
; Publication No. US2003027998A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-227-999
; CURRENT APPLICATION NUMBER: US/09/796,753
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 09/183,175
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 09/223,094
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/223,546
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/224,246
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/259,388
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 60/122,458
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 09/312,359
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/336,536
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: 09/342,687
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 09/345,464
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: 09/365,164
```

```
        ; PRIOR FILING DATE: 1999-07-30
        ; PRIOR APPLICATION NUMBER: 09/399,723
        ; PRIOR FILING DATE: 1999-09-20
        ; PRIOR APPLICATION NUMBER: 09/409,634
        ; PRIOR FILING DATE: 1999-09-30
        ; PRIOR APPLICATION NUMBER: 09/471,179
        ; PRIOR FILING DATE: 1999-12-23
        ; PRIOR APPLICATION NUMBER: 09/474,071
        ; PRIOR FILING DATE: 1999-12-29
        ; PRIOR APPLICATION NUMBER: 09/474,072
        ; PRIOR FILING DATE: 1999-12-29
        ; PRIOR APPLICATION NUMBER: 09/514,010
        ; PRIOR FILING DATE: 2000-02-25
        ; PRIOR APPLICATION NUMBER: 09/516,745
        ; PRIOR FILING DATE: 2000-03-01
        ; PRIOR APPLICATION NUMBER: 09/572,002
        ; PRIOR FILING DATE: 2000-05-14
        ; PRIOR APPLICATION NUMBER: 09/597,993
        ; PRIOR FILING DATE: 2000-06-19
        ; PRIOR APPLICATION NUMBER: 09/599,596
        ; PRIOR FILING DATE: 2000-06-22
        ; PRIOR APPLICATION NUMBER: 09/630,334
        ; PRIOR FILING DATE: 2000-07-31
        ; PRIOR APPLICATION NUMBER: 09/606,565
        ; PRIOR FILING DATE: 2000-06-29
        ; PRIOR APPLICATION NUMBER: 09/606,317
        ; PRIOR FILING DATE: 2000-06-29
        ; PRIOR APPLICATION NUMBER: 09/665,666
        ; PRIOR FILING DATE: 2000-09-20
        ; PRIOR APPLICATION NUMBER: 09/677,751
        ; PRIOR FILING DATE: 2000-09-30
        ; NUMBER OF SEQ ID NOS: 162
        ; SEQ ID NO 12
        ; LENGTH: 333
        ; TYPE: PRT
        ; ORGANISM: Homo sapiens
US-09-796-753-12

Query Match      100.0%; Score 899; DB 10; Length 333;
Best Local Similarity 100.0%; Pred. No. 1e-86;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 DLYFIIDKSGSVLHNNHNIYFYFVQLAHKFIISPOLRMSFVIFSTRGTTLMKLTEDRQIR 60
        |||||||
DB      44 DLYFIIDKSGSVLHNNHNIYFYFVQLAHKFIISPOLRMSFVIFSTRGTTLMKLTEDRQIR 103
        |||||||

QY      61 QGLEELQKVLPGSDTYMHGFEFASQIYYENRGYRTASVITALTDEGLHEDLFFYSER 120
        |||||||
DB      104 QGLEELQKVLPGSDTYMHGFEFASQIYYENRGYRTASVITALTDEGLHEDLFFYSER 163
        |||||||

QY      121 EANRSDLGAIYCVGVKDFNETOLARIADSKDHVFPVNDGFQALOGIHSIL 173
        |||||||
DB      164 EANRSDLGAIYCVGVKDFNETOLARIADSKDHVFPVNDGFQALOGIHSIL 216
        |||||||

RESULT 4
US-10-038-307-2
; Sequence 2, Application US/10038307
; Publication No. US20030134786A1
; GENERAL INFORMATION:
; APPLICANT: James B. ROTTMAN
; APPLICANT: Theresa L. O'KEEFE
; APPLICANT: Engin OZKAYNAK
; APPLICANT: Judith J. HEALEY
; TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
; FILE REFERENCE: 7853-253-999
; CURRENT APPLICATION NUMBER: US/10/038,307
; CURRENT FILING DATE: 2002-06-28
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 333
; TYPE: PRT
```

ORGANISM: Homo sapiens
US-10-038-307-2

Query Match 100.0%; Score 899; DB 14; Length 333;
Best Local Similarity 100.0%; Pred. No. 1e-86;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLYFIIDKSGSVLHHNNEIYFVEQLAHKFIISPOLRMSFIVFSTRGTTLMKLTEDRQIR 60
DB 44 DLYFIIDKSGSVLHHNNEIYFVEQLAHKFIISPOLRMSFIVFSTRGTTLMKLTEDRQIR 103
QY 61 OGLEELQKVLPGDDTYMHGFEFASBQIYYENRQGYRTASVIALTDGELHEDLFFYSER 120
DB 104 OGLEELQKVLPGDDTYMHGFEFASBQIYYENRQGYRTASVIALTDGELHEDLFFYSER 163
QY 121 EANRSRDLGAIYCVGVKDFNETOLARIADSKDHFVPVNDGFOALQGIHSIL 173
DB 164 EANRSRDLGAIYCVGVKDFNETOLARIADSKDHFVPVNDGFOALQGIHSIL 216

RESULT 5

US-10-201-292-2
Sequence 2, Application US/10201292
Publication No. US20030144193A1

GENERAL INFORMATION:
APPLICANT: James B. ROTTMAN
APPLICANT: Theresa L. O'KEEFE
APPLICANT: Engin OZKAYNAK
APPLICANT: Judith J. HEALEY
TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
FILE REFERENCE: 7853-253-999
CURRENT APPLICATION NUMBER: US/10/201,292
TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
FILE REFERENCE: 7853-253-999
CURRENT FILING DATE: 2003-02-14
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 333
TYPE: PRT
ORGANISM: Homo sapiens
US-10-201-292-2

Query Match 100.0%; Score 899; DB 14; Length 333;
Best Local Similarity 100.0%; Pred. No. 1e-86;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLYFIIDKSGSVLHHNNEIYFVEQLAHKFIISPOLRMSFIVFSTRGTTLMKLTEDRQIR 60
DB 44 DLYFIIDKSGSVLHHNNEIYFVEQLAHKFIISPOLRMSFIVFSTRGTTLMKLTEDRQIR 103
QY 61 OGLEELQKVLPGDDTYMHGFEFASBQIYYENRQGYRTASVIALTDGELHEDLFFYSER 120
DB 104 OGLEELQKVLPGDDTYMHGFEFASBQIYYENRQGYRTASVIALTDGELHEDLFFYSER 163
QY 121 EANRSRDLGAIYCVGVKDFNETOLARIADSKDHFVPVNDGFOALQGIHSIL 173
DB 164 EANRSRDLGAIYCVGVKDFNETOLARIADSKDHFVPVNDGFOALQGIHSIL 216

RESULT 6

US-10-038-307-22
Sequence 22, Application US/10038307
Publication No. US20030134786A1

GENERAL INFORMATION:
APPLICANT: James B. ROTTMAN
APPLICANT: Theresa L. O'KEEFE
APPLICANT: Engin OZKAYNAK
APPLICANT: Judith J. HEALEY
TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
FILE REFERENCE: 7853-253-999
CURRENT APPLICATION NUMBER: US/10/038,307
CURRENT FILING DATE: 2002-06-28
NUMBER OF SEQ ID NOS: 26
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 22
LENGTH: 342
TYPE: PRT
ORGANISM: Homo sapiens
US-10-038-307-22

Query Match 100.0%; Score 899; DB 14; Length 342;
Best Local Similarity 100.0%; Pred. No. 1.1e-86;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLYFIIDKSGSVLHHNNEIYFVEQLAHKFIISPOLRMSFIVFSTRGTTLMKLTEDRQIR 60
DB 53 DLYFIIDKSGSVLHHNNEIYFVEQLAHKFIISPOLRMSFIVFSTRGTTLMKLTEDRQIR 112
QY 61 OGLEELQKVLPGDDTYMHGFEFASBQIYYENRQGYRTASVIALTDGELHEDLFFYSER 120
DB 113 OGLEELQKVLPGDDTYMHGFEFASBQIYYENRQGYRTASVIALTDGELHEDLFFYSER 172
QY 121 EANRSRDLGAIYCVGVKDFNETOLARIADSKDHFVPVNDGFOALQGIHSIL 173
DB 173 EANRSRDLGAIYCVGVKDFNETOLARIADSKDHFVPVNDGFOALQGIHSIL 225

RESULT 7

US-10-201-292-22
Sequence 22, Application US/10201292
Publication No. US20030144193A1

GENERAL INFORMATION:
APPLICANT: James B. ROTTMAN
APPLICANT: Theresa L. O'KEEFE
APPLICANT: Engin OZKAYNAK
APPLICANT: Judith J. HEALEY
TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
FILE REFERENCE: 7853-253-999
CURRENT APPLICATION NUMBER: US/10/201,292
TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
FILE REFERENCE: 7853-253-999
CURRENT FILING DATE: 2003-02-14
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 22
LENGTH: 342
TYPE: PRT
ORGANISM: Homo sapiens
US-10-201-292-22

Query Match 100.0%; Score 899; DB 14; Length 342;
Best Local Similarity 100.0%; Pred. No. 1.1e-86;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLYFIIDKSGSVLHHNNEIYFVEQLAHKFIISPOLRMSFIVFSTRGTTLMKLTEDRQIR 60
DB 53 DLYFIIDKSGSVLHHNNEIYFVEQLAHKFIISPOLRMSFIVFSTRGTTLMKLTEDRQIR 112
QY 61 OGLEELQKVLPGDDTYMHGFEFASBQIYYENRQGYRTASVIALTDGELHEDLFFYSER 120
DB 113 OGLEELQKVLPGDDTYMHGFEFASBQIYYENRQGYRTASVIALTDGELHEDLFFYSER 172
QY 121 EANRSRDLGAIYCVGVKDFNETOLARIADSKDHFVPVNDGFOALQGIHSIL 173
DB 173 EANRSRDLGAIYCVGVKDFNETOLARIADSKDHFVPVNDGFOALQGIHSIL 225

RESULT 8

US-10-038-307-24
Sequence 24, Application US/10038307
Publication No. US20030134786A1

GENERAL INFORMATION:
APPLICANT: James B. ROTTMAN
APPLICANT: Theresa L. O'KEEFE
APPLICANT: Engin OZKAYNAK
APPLICANT: Judith J. HEALEY
TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
FILE REFERENCE: 7853-253-999
CURRENT APPLICATION NUMBER: US/10/038,307

;; CURRENT FILING DATE: 2002-06-28
;; NUMBER OF SEQ ID NOS: 26
;; SOFTWARE: FASTSEQ for Windows Version 4.0
;; SEQ ID NO 24
;; LENGTH: 345
;; TYPE: PRF
;; ORGANISM: Homo sapiens
US-10-038-307-24

Query Match 100.0%; Score 899; DB 14; Length 345;
Best Local Similarity 100.0%; Pred. No. 1.1e-86;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLYFIIDKSGSVLHNNHNIYFVEQLAHKFIISPOLRMSFIVFSTRTGTTLMKLTEDREQIR 60
DB 44 DLYFIIDKSGSVLHNNHNIYFVEQLAHKFIISPOLRMSFIVFSTRTGTTLMKLTEDREQIR 103
QY 61 OGLEELQKVLPGGDTYMHGFEFASBOIYYENRQGYRTASVIALTDGELHEDLFFYSER 120
DB 104 OGLEELQKVLPGGDTYMHGFEFASBOIYYENRQGYRTASVIALTDGELHEDLFFYSER 163
QY 121 EANRSRDLGAIYVCVGVKDFNETQLARIADSKDHVPVNDGFQALQGIHSIL 173
DB 164 EANRSRDLGAIYVCVGVKDFNETQLARIADSKDHVPVNDGFQALQGIHSIL 216

RESULT 9

US-10-201-292-24
;; Sequence 24, Application US/10201292
;; Publication No. US20030144193A1
;; GENERAL INFORMATION:
;; APPLICANT: James B. ROTTMAN
;; APPLICANT: Theresa L. O'KEEFE
;; APPLICANT: Engin OZKAYNAK
;; APPLICANT: Judith J. HEALEY
;; TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
;; FILE REFERENCE: 7853-253-999
;; CURRENT APPLICATION NUMBER: US/10/201,292
;; CURRENT FILING DATE: 2003-02-14
;; NUMBER OF SEQ ID NOS: 36
;; SOFTWARE: FASTSEQ for Windows Version 4.0
;; SEQ ID NO 24
;; LENGTH: 345
;; TYPE: PRF
;; ORGANISM: Homo sapiens
US-10-201-292-24

Query Match 100.0%; Score 899; DB 14; Length 345;
Best Local Similarity 100.0%; Pred. No. 1.1e-86;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLYFIIDKSGSVLHNNHNIYFVEQLAHKFIISPOLRMSFIVFSTRTGTTLMKLTEDREQIR 60
DB 44 DLYFIIDKSGSVLHNNHNIYFVEQLAHKFIISPOLRMSFIVFSTRTGTTLMKLTEDREQIR 103
QY 61 OGLEELQKVLPGGDTYMHGFEFASBOIYYENRQGYRTASVIALTDGELHEDLFFYSER 120
DB 104 OGLEELQKVLPGGDTYMHGFEFASBOIYYENRQGYRTASVIALTDGELHEDLFFYSER 163
QY 121 EANRSRDLGAIYVCVGVKDFNETQLARIADSKDHVPVNDGFQALQGIHSIL 173
DB 164 EANRSRDLGAIYVCVGVKDFNETQLARIADSKDHVPVNDGFQALQGIHSIL 216

RESULT 10

US-09-833-245-621
;; Sequence 621, Application US/09833245
;; Publication No. US20040010134A1
;; GENERAL INFORMATION:
;; APPLICANT: Human Genome Sciences, Inc.
;; TITLE OF INVENTION: Albumin Fusion Proteins
;; FILE REFERENCE: PF546PCT
;; CURRENT APPLICATION NUMBER: US/09/833,245

;; CURRENT FILING DATE: 2001-04-12
;; PRIOR APPLICATION NUMBER: 60/229, 358
;; PRIOR FILING DATE: 2000-04-12
;; PRIOR APPLICATION NUMBER: 60/256, 931
;; PRIOR FILING DATE: 2000-12-21
;; PRIOR APPLICATION NUMBER: 60/199, 384
;; PRIOR FILING DATE: 2000-04-25
;; NUMBER OF SEQ ID NOS: 2267
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 621
;; LENGTH: 403
;; TYPE: PRF
;; ORGANISM: Homo sapiens
US-09-833-245-621

Query Match 100.0%; Score 899; DB 11; Length 403;
Best Local Similarity 100.0%; Pred. No. 1.4e-86;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLYFIIDKSGSVLHNNHNIYFVEQLAHKFIISPOLRMSFIVFSTRTGTTLMKLTEDREQIR 60
DB 44 DLYFIIDKSGSVLHNNHNIYFVEQLAHKFIISPOLRMSFIVFSTRTGTTLMKLTEDREQIR 103
QY 61 OGLEELQKVLPGGDTYMHGFEFASBOIYYENRQGYRTASVIALTDGELHEDLFFYSER 120
DB 104 OGLEELQKVLPGGDTYMHGFEFASBOIYYENRQGYRTASVIALTDGELHEDLFFYSER 163
QY 121 EANRSRDLGAIYVCVGVKDFNETQLARIADSKDHVPVNDGFQALQGIHSIL 173
DB 164 EANRSRDLGAIYVCVGVKDFNETQLARIADSKDHVPVNDGFQALQGIHSIL 216

RESULT 11

US-10-201-292-28
;; Sequence 28, Application US/10201292
;; Publication No. US20030144193A1
;; GENERAL INFORMATION:
;; APPLICANT: James B. ROTTMAN
;; APPLICANT: Theresa L. O'KEEFE
;; APPLICANT: Engin OZKAYNAK
;; APPLICANT: Judith J. HEALEY
;; TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
;; FILE REFERENCE: 7853-253-999
;; CURRENT APPLICATION NUMBER: US/10/201,292
;; CURRENT FILING DATE: 2003-02-14
;; NUMBER OF SEQ ID NOS: 36
;; SOFTWARE: FASTSEQ for Windows Version 4.0
;; SEQ ID NO 28
;; LENGTH: 460
;; TYPE: PRF
;; ORGANISM: Homo sapiens
US-10-201-292-28

Query Match 100.0%; Score 899; DB 14; Length 460;
Best Local Similarity 100.0%; Pred. No. 1.6e-86;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLYFIIDKSGSVLHNNHNIYFVEQLAHKFIISPOLRMSFIVFSTRTGTTLMKLTEDREQIR 60
DB 44 DLYFIIDKSGSVLHNNHNIYFVEQLAHKFIISPOLRMSFIVFSTRTGTTLMKLTEDREQIR 103
QY 61 OGLEELQKVLPGGDTYMHGFEFASBOIYYENRQGYRTASVIALTDGELHEDLFFYSER 120
DB 104 OGLEELQKVLPGGDTYMHGFEFASBOIYYENRQGYRTASVIALTDGELHEDLFFYSER 163
QY 121 EANRSRDLGAIYVCVGVKDFNETQLARIADSKDHVPVNDGFQALQGIHSIL 173
DB 164 EANRSRDLGAIYVCVGVKDFNETQLARIADSKDHVPVNDGFQALQGIHSIL 216

RESULT 12

US-10-201-292-30
;; Sequence 30, Application US/10201292

```
/ Publication NO. US20030144193A1
/ GENERAL INFORMATION:
/ APPLICANT: James B. ROTTMAN
/ APPLICANT: Theresa L. O'KEEFE
/ APPLICANT: Engin OZKAYNAK
/ APPLICANT: Judith J. HEALEY
/ TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
/ FILE REFERENCE: 7853-253-999
/ CURRENT APPLICATION NUMBER: US/10/201.292
/ CURRENT FILING DATE: 2003-02-14
/ NUMBER OF SEQ ID NOS: 36
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 30
/ LENGTH: 460
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-201-292-30
```

```
Query Match          100.0%; Score 899; DB 14; Length 460;
Best Local Similarity 100.0%; Pred. No. 1.6e-86;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 DLYFILDKSGSVLHNMNEIYFVEQLAHKFI SPQLRMSFIVFSTRGTTMLKLTEDRQIR 60
Db 44 DLYFILDKSGSVLHNMNEIYFVEQLAHKFI SPQLRMSFIVFSTRGTTMLKLTEDRQIR 103
Qy 61 OGLEELQKVLPGGDTYHHEGFERASEQIYYENRGYRTASVIALTDGELHEDLFFYSER 120
Db 104 OGLEELQKVLPGGDTYHHEGFERASEQIYYENRGYRTASVIALTDGELHEDLFFYSER 163
Qy 121 EANRSRDIGAIIVYCVGVKDFNETOLARIADSKDHVFPVNDGFOALQGIHSIL 173
Db 164 EANRSRDIGAIIVYCVGVKDFNETOLARIADSKDHVFPVNDGFOALQGIHSIL 216
```

```
RESULT 13
US-10-201-292-32
/ Sequence 32, Application US/10201292
/ Publication No. US20030144193A1
/ GENERAL INFORMATION:
/ APPLICANT: James B. ROTTMAN
/ APPLICANT: Theresa L. O'KEEFE
/ APPLICANT: Engin OZKAYNAK
/ APPLICANT: Judith J. HEALEY
/ TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
/ FILE REFERENCE: 7853-253-999
/ CURRENT APPLICATION NUMBER: US/10/201.292
/ CURRENT FILING DATE: 2003-02-14
/ NUMBER OF SEQ ID NOS: 36
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 32
/ LENGTH: 479
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-201-292-32
```

```
Query Match          100.0%; Score 899; DB 14; Length 479;
Best Local Similarity 100.0%; Pred. No. 1.7e-86;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 DLYFILDKSGSVLHNMNEIYFVEQLAHKFI SPQLRMSFIVFSTRGTTMLKLTEDRQIR 60
Db 44 DLYFILDKSGSVLHNMNEIYFVEQLAHKFI SPQLRMSFIVFSTRGTTMLKLTEDRQIR 103
Qy 61 OGLEELQKVLPGGDTYHHEGFERASEQIYYENRGYRTASVIALTDGELHEDLFFYSER 120
Db 104 OGLEELQKVLPGGDTYHHEGFERASEQIYYENRGYRTASVIALTDGELHEDLFFYSER 163
Qy 121 EANRSRDIGAIIVYCVGVKDFNETOLARIADSKDHVFPVNDGFOALQGIHSIL 173
Db 164 EANRSRDIGAIIVYCVGVKDFNETOLARIADSKDHVFPVNDGFOALQGIHSIL 216
```

```
RESULT 14
US-10-201-292-34
/ Sequence 34, Application US/10201292
/ Publication No. US20030144193A1
/ GENERAL INFORMATION:
/ APPLICANT: James B. ROTTMAN
/ APPLICANT: Theresa L. O'KEEFE
/ APPLICANT: Engin OZKAYNAK
/ APPLICANT: Judith J. HEALEY
/ TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
/ FILE REFERENCE: 7853-253-999
/ CURRENT APPLICATION NUMBER: US/10/201.292
/ CURRENT FILING DATE: 2003-02-14
/ NUMBER OF SEQ ID NOS: 36
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 34
/ LENGTH: 504
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-201-292-34
```

```
Query Match          100.0%; Score 899; DB 14; Length 504;
Best Local Similarity 100.0%; Pred. No. 1.8e-86;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 DLYFILDKSGSVLHNMNEIYFVEQLAHKFI SPQLRMSFIVFSTRGTTMLKLTEDRQIR 60
Db 44 DLYFILDKSGSVLHNMNEIYFVEQLAHKFI SPQLRMSFIVFSTRGTTMLKLTEDRQIR 103
Qy 61 OGLEELQKVLPGGDTYHHEGFERASEQIYYENRGYRTASVIALTDGELHEDLFFYSER 120
Db 104 OGLEELQKVLPGGDTYHHEGFERASEQIYYENRGYRTASVIALTDGELHEDLFFYSER 163
Qy 121 EANRSRDIGAIIVYCVGVKDFNETOLARIADSKDHVFPVNDGFOALQGIHSIL 173
Db 164 EANRSRDIGAIIVYCVGVKDFNETOLARIADSKDHVFPVNDGFOALQGIHSIL 216
```

```
RESULT 15
US-10-201-292-36
/ Sequence 36, Application US/10201292
/ Publication No. US20030144193A1
/ GENERAL INFORMATION:
/ APPLICANT: James B. ROTTMAN
/ APPLICANT: Theresa L. O'KEEFE
/ APPLICANT: Engin OZKAYNAK
/ APPLICANT: Judith J. HEALEY
/ TITLE OF INVENTION: Tango 197 and Tango 216 Compositions and Methods
/ FILE REFERENCE: 7853-253-999
/ CURRENT APPLICATION NUMBER: US/10/201.292
/ CURRENT FILING DATE: 2003-02-14
/ NUMBER OF SEQ ID NOS: 36
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 36
/ LENGTH: 529
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-201-292-36
```

```
Query Match          100.0%; Score 899; DB 14; Length 529;
Best Local Similarity 100.0%; Pred. No. 2e-86;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 DLYFILDKSGSVLHNMNEIYFVEQLAHKFI SPQLRMSFIVFSTRGTTMLKLTEDRQIR 60
Db 44 DLYFILDKSGSVLHNMNEIYFVEQLAHKFI SPQLRMSFIVFSTRGTTMLKLTEDRQIR 103
Qy 61 OGLEELQKVLPGGDTYHHEGFERASEQIYYENRGYRTASVIALTDGELHEDLFFYSER 120
Db 104 OGLEELQKVLPGGDTYHHEGFERASEQIYYENRGYRTASVIALTDGELHEDLFFYSER 163
Qy 121 EANRSRDIGAIIVYCVGVKDFNETOLARIADSKDHVFPVNDGFOALQGIHSIL 173
```

Db 164 EARNRDLGAIYCVGVKDFNFTQLARIADSKDHVPVNDGFOALOGIIHSIL 216

Search completed: June 13, 2005, 20:36:51
Job time : 56.4703 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 19:47:08 ; Search time 12.4119 Seconds
(without alignments)
1341.095 Million cell updates/sec

Title: US-09-970-076-2_COPY_44_216

Perfect score: 899
Sequence: 1 DLVFIIDKSGSVLHMHNEIY.....HFPVNDGFOALGQIHSIL 173

Scoring table: BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129.5	14.4	724	2 A48569	antigen Em100 - B1
2	128.5	14.3	1153	2 RWHUB	cell surface glyco
3	125	13.9	3124	2 A40020	collagen alpha 1(X
4	121.5	13.5	712	2 A45638	immunoglobulin mic
5	121.5	13.5	1163	1 RWHUIC	cell surface glyco
6	120.5	13.4	1153	2 S00551	leukocyte surface
7	113.5	12.6	760	1 C2MS	classical-compleme
8	110	12.2	1747	2 A45974	collagen alpha 1(X
9	110	12.2	1857	2 S31212	collagen alpha 1(X
10	110	12.2	1888	2 S78476	collagen alpha 1(X
11	109	12.1	3051	2 S42373	hypothetical prote
12	104.5	11.6	929	2 I51027	type XII collagen
13	102	11.3	2944	2 A54849	collagen alpha 1(V
14	101.5	11.3	764	2 BBHU	complement factor
15	100.5	11.2	567	2 T28797	hypothetical prote
16	100	11.1	460	2 T23087	hypothetical prote
17	99	11.0	698	2 B96958	dnak protein (heat
18	98.5	11.0	1029	1 S21369	collagen alpha 2(V
19	96.5	10.7	761	1 BBMS	complement factor
20	93	10.3	3176	2 CGHJ7A	collagen alpha 3(V
21	92.5	10.3	3137	2 A37797	collagen alpha 3(V
22	91.5	10.2	1179	2 A53213	integrin alpha-B c
23	88.5	9.8	932	2 JCS953	inter-alpha-inhibi
24	88	9.8	917	2 S09646	collagen alpha 2(V
25	88	9.8	1018	1 CGHJ2A	collagen alpha 2(V
26	88	9.8	1318	2 T30889	valine-tRNA ligase
27	87.5	9.7	341	2 T32989	hypothetical prote
28	87.5	9.7	418	2 AG2350	hypothetical prote
29	87	9.7	710	2 AG2285	hypothetical prote

30	86.5	9.6	741	2 T46488	hypothetical prote
31	85.5	9.5	843	2 A40970	undulin 1 - human
32	84.5	9.4	800	2 AB1079	endoglucanase homo
33	84.5	9.4	875	1 ITRECAP	DNA topoisomerase
34	84.5	9.4	875	2 B91018	DNA gyrase subunit
35	84.5	9.4	875	2 D65862	DNA gyrase, subunit
36	84	9.3	689	2 F84811	probable retroelem
37	84	9.3	2954	2 T14156	kinesin-related pr
38	83.5	9.3	427	2 AH2350	hypothetical prote
39	82.5	9.2	340	2 B83262	hypothetical prote
40	82.5	9.2	493	2 A33809	cartilage matrix p
41	82	9.1	487	1 VZEBPT	sensor kinase phoQ
42	81.5	9.1	340	2 E70121	hypothetical prote
43	81	9.0	5369	2 T44807	mycosubtilin synth
44	80.5	9.0	448	2 B71257	conserved hypothet
45	80.5	9.0	673	2 JC2458	dnak-type molecula

ALIGNMENTS

RESULT 1
A48569
antigen Em100 - Eimeria maxima
C/Species: Eimeria maxima
C/Date: 01-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C/Accession: A48569
R:Pasamonies, L.; Hug, D.; Humbelin, M.; Weber, G.
Mol. Biochem. Parasitol. 57, 171-174, 1993
A/Title: Sequence of a major Eimeria maxima antigen homologous to the Eimeria tenella m
A/Reference number: A48569; MUID:93149203; PMID:8426611
A/Accession: A48569
A/Status: preliminary
A/Molecule type: nucleic acid
A/Residues: 1-724 <PAS>
A/Cross-References: UNIPROT:Q04588; GB:M99058; NID:g158890; PID:g158891
A/Note: Sequence extracted from NCBI backbone (NCBI:123776, NCBI:123777)
F/45-218/Domain: von Willebrand factor type A repeat homology <WAA2>
F/238-239/Domain: thrombospondin type 1 repeat homology <THR1>
F/309-371/Domain: thrombospondin type 1 repeat homology <THR2>
F/372-432/Domain: thrombospondin type 1 repeat homology <THR3>
F/433-493/Domain: thrombospondin type 1 repeat homology <THR4>
F/494-556/Domain: thrombospondin type 1 repeat homology <THR5>
F/560-610/Domain: thrombospondin type 1 repeat homology <THR6>

Query Match 14.4%; Score 129.5; DB 2; Length 724;
Best Local Similarity 28.1%; Pred. No. 0.002;
Matches 48; Conservative 30; Mismatches 68; Indels 25; Gaps 10;
Qy 1 DLVFIIDKSGSV-LHMHNEIYFVEQLAHKF-ISP-QLRMSFIVSTRTLMKLTEDRE 57
Db DVLMLVDESSISGNSNGKRSFISNPAQTMPLSPDVRGVIVFGISATYRMLSSRA 106
Qy 58 QIRGIEELQKLP-~~CGDTYMEGFERASBQIYENRGYR--TASVIALTDGLHED~~ 113
Db 107 QNADLLAAAKKL~~FVAAGSYTHLGLAKA--BEILFSGKGRDVA~~PKMILMTDGA---- 161
Qy 114 LPTFSERE-----~~ANRSRLGALVYCVGV-KDENEQLARIA--DKDNV~~ 155
Db 162 -----SSRSQTL~~SAAEKLRNRGVITIVLVGTVGVSACRSIACDTS~~DTV 208

RESULT 2

RWHUB
cell surface glycoprotein CD11b precursor [validated] - human
N/Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein Ma
eukocyte integrin alpha chain; neutrophil adherence receptor alpha chain
C/Species: Homo sapiens (man)
C/Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C/Accession: A31108; A28915; A31600; A30892; A46536; A26091; I52567
R/Corby, A.L.; Kishimoto, T.K.; Miller, U.U.; Springer, T.A.
J. Biol. Chem. 263, 12403-12411, 1988
A/Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11

B.
 A:Reference number: A31108; MUID:88315033; PMID:2457584
 A:Accession: A31108
 A:Molecule type: mRNA
 A:Residues: 1-1153 <COR>
 A:Cross-references: UNIPROT:P11215; GB:J03925; NID:G187284; PIDN:AAA59544.1; PID:G307148
 A>Note: part of this sequence was confirmed by protein sequencing
 R:Arnaut, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.
 J. Cell Biol. 106, 2153-2158, 1988
 A:Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor M
 A:Reference number: A28915; MUID:88257215; PMID:2454931
 A:Accession: A28915
 A:Molecule type: mRNA
 A:Residues: 1-499, 501-965, 'P', 967-1153 <ARN>
 A:Cross-references: GB:M18044; GB:U03270; GB:M19664; GB:X07421; NID:G186935; PIDN:AAA594
 A>Note: the authors translated the codon TAC for residue 1129 as Thr
 A:Note: part of this sequence, including the amino end of the mature protein, was confir
 R:Shelley, C.S.; Arnaut, M.A.
 Proc. Natl. Acad. Sci. U.S.A. 85, 10525-10529, 1988
 A:Title: The promoter of the CD11b gene directs myeloid-specific and developmentally reg
 A:Reference number: A41600; MUID:92073318; PMID:1683702
 A:Accession: A41600
 A:Molecule type: DNA
 A:Residues: 1-9 <SHE>
 A:Cross-references: GB:M76724; NID:G180018; PIDN:AAA58410.1; PID:G553215
 R:Arnaut, M.A.; Remold-O'Donnell, B.; Pierce, M.W.; Harris, P.; Tenen, D.G.
 Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988
 A:Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhesion
 A:Reference number: A94193; MUID:88190151; PMID:2833753
 A:Accession: A30892
 A:Molecule type: mRNA
 A:Residues: 917-1042 <AR2>
 A:Cross-references: GB:M18044
 R:Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.W.; Back, A.L.; Roth, G.J.
 Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989
 A:Title: cDNA sequence for the alpha subunit of the human neutrophil adherence receptor
 A:Reference number: A32218; MUID:89098893; PMID:2553162
 A:Accession: A32218
 A:Molecule type: mRNA
 A:Residues: 9-1153 <HIC>
 A:Cross-references: GB:U04145; NID:G189068; PIDN:AAA59903.1; PID:G386975
 A:Note: part of this sequence was confirmed by protein sequencing
 R:Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.
 J. Immunol. 150, 480-490, 1993
 A:Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-
 n during evolution.
 A:Reference number: A46526; MUID:93123748; PMID:8419480
 A:Accession: A46526
 A:Molecule type: DNA
 A:Status: not compared with conceptual translation
 A:Residues: 1-499, 501-1153 <FLB>
 A:Cross-references: GB:S52227; NID:G263047; PIDN:AA24821.1; PID:G263049
 A:Note: the last three bases of intron 13, CAG, are included in some but not all mature
 A:Note: sequence extracted from NCBI backbone (NCBI:P:121963)
 R:Pierce, M.W.; Remold-O'Donnell, B.; Todd III, R.F.; Arnaut, M.A.
 Biochim. Biophys. Acta 874, 368-371, 1986
 A:Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across sp
 A:Reference number: A90664; MUID:87076671; PMID:35359202
 A:Accession: A26091
 A:Molecule type: protein
 A:Residues: 17-31 <PIE>
 A:Experimental source: granulocytes
 R:Pahl, H.L.; Rossmartin, A.G.; Tenen, D.G.
 Blood 79, 865-870, 1992
 A:Title: Characterization of the myeloid-specific CD11b promoter.
 A:Reference number: 152567; MUID:92144986; PMID:1346576
 A:Accession: 152567
 A:Status: translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-9 <RES>
 A:Cross-references: GB:M84477; NID:G180184; PIDN:AAA51960.1; PID:G553219
 C:Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1
 C:Genetic:

A:Gene: GDB:ITGAM, CR3A
 A:Cross-references: GDB:120599; OMIM:120980
 A:Map position: 16p11.2-16p11.2
 A:Note: promoter contains a GATA motif and two Sp1 consensus binding sites
 C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homo
 C:Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; magr
 F:1-16/Domain: signal sequence #status predicted <SIG>
 F:17-1153/Product: cell surface glycoprotein CD11b #status experimental <MAT>
 F:17-1153/Domain: extracellular #status predicted <EXT>
 F:148-318/Domain: von Willebrand factor type A repeat homology <VWA2>
 F:465-473/Region: calcium/magnesium binding #status predicted
 F:530-538/Region: calcium/magnesium binding #status predicted
 F:593-601/Region: calcium/magnesium binding #status predicted
 F:1109-1134/Domain: transmembrane #status predicted <TM>
 F:1135-1153/Domain: intracellular #status predicted <INT>
 F:86,240,391,469,693,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Binding
 Query Match 14.3%; Score 128.5; DB 1; Length 1153;
 Best Local Similarity 26.7%; Pred. No. 0.0042; Indels 39; Gaps 11;
 Matches 51; Conservative 38; Mismatches 63;
 QY 1 DLYFIIDKSGSVL-HHMEIYFVEQLAHKFIISPOLMSFIVSTGTTMLKTED---- 55
 Db 150 DIAFLIDGSGSIIHPHFRMKFVST-----VMEQLKSKTLFS-----LMQYSEERRH 129
 QY 56 -----RQIQGLEBLQKLPVGGDTYWHSEFPAESQIYYENRGYRTA-SVITALT 106
 Db 200 FTFKERNPNPNRSLVKVPIYQLT--GRITVATGIRKVRRELFTITGARKNAKFLIVIT 257
 QY 107 DGLHEDLPFYSE--REANSRDLGALVYCVGKDNENQTLAR-----INDS--KDHVFP 157
 Db 258 DGEKFDPLGLEYEVITREADRE--GVIRYIVGDAFRSEKSHQELNTTISKPRDHVFO 314
 QY 158 VNDGFQALQGI 168
 Db 315 VNN-PEALKTI 324
 RESULT 3
 A40020
 collagen alpha 1(XII) chain precursor - chicken
 N:Alternate names: fibronectin
 C:Species: Gallus gallus (chicken)
 C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #ext change 09-Jul-2004
 C:Accession: A40020; A34485; B34485; A28037; S23814; S22524; S28811
 R:Yamagata, M.; Yamada, K.M.; Yamada, S.S.; Shinomura, T.; Tanaka, H.; Nishida, Y.; Obari
 J. Cell Biol. 115, 209-221, 1991
 A:Title: The complete primary structure of type XII collagen shows a chimeric molecule w
 nous region of type IX collagen, and short collagenous domains with an Arg-Gly-Asp site.
 A:Reference number: A40020; MUID:92011862; PMID:1918137
 A:Accession: A40020
 A:Molecule type: mRNA
 A:Residues: 1-3124 <YAM>
 A:Cross-references: UNIPROT:P13944; GB:D00824; NID:G222810; PIDN:BA40070.1; PID:G222811
 A:Note: in the authors' translation residues 1216-1219 are shown after residue 1235 and,
 R:Gordon, M.K.; Gerecke, D.R.; Duddle, B.; van der Reet, M.; Olsen, B.R.
 J. Biol. Chem. 264, 19772-19778, 1989
 A:Title: Type XII collagen. A large multidomain molecule with partial homology to type I
 A:Reference number: A34485; MUID:90062079; PMID:2584192
 A:Accession: A34485
 A:Molecule type: mRNA
 A:Residues: 2456-2758, 'A', 2760-2802, 'F', 2804-2976, 'F', 2978-3124 <GOR>
 A:Cross-references: EMBL:U05137; NID:G211284; PIDN:AAA48635.1; PID:G211285
 A:Accession: B34485
 A:Molecule type: protein
 A:Residues: 2772-2792; 2846-2873 <GOR2>
 R:Gordon, M.K.; Gerecke, D.R.; Olsen, B.R.
 Proc. Natl. Acad. Sci. U.S.A. 84, 6040-6044, 1987
 A:Title: Type XII collagen: distinct extracellular matrix component discovered by cDNA
 A:Reference number: A28037; MUID:87317590; PMID:3476925
 A:Accession: A28037
 A:Molecule type: mRNA
 A:Residues: 2960-2976, 'F', 2978-3074, 'AG' <GOR3>

A/Cross-references: EMBL:MI7375; NID:G211649; PIDN:AAA8718.1; PID:G211650
A/Note: this sequence has been revised in reference A34485
Eur. J. Biochem. 207, 847-856, 1992
A/Title: A major oligomeric fibroblast proteoglycan identified as a novel large form of
A/Reference number: 523814; MUID:92362621; PMID:1332460
A/Accession: 523814
A/Molecule type: protein
A/Residues: 'X',1333,'Q',1335-1347,1514-1928;2504,'X',2506,'X',2508-2511,'X',2513-2517 <DDB>
R/Db/let: B., van der Ree, M.
J. Biol. Chem. 262, 17724-17727, 1987
A/Title: Type XII collagen is expressed in embryonic chick tendons. Isolation of pepsin-
A/Reference number: 522254; MUID:88087065; PMID:3121603
A/Accession: 522254
A/Molecule type: protein
A/Residues: 2831-2832,'T',2834,'R',2836-2843;3002-3014 <DDB>
R/Trieb, J., Trueb, B.
Biochim. Biophys. Acta 1171, 97-98, 1992
A/Title: The two splice variants of collagen XII share a common 5' end.
A/Reference number: 528811; MUID:93042014; PMID:1420368
A/Accession: 528811
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-24,1189-1257,'S',1259-1263,'E',1265-1280 <TRU>
A/Cross-references: EMBL:X67327
A/Note: B.
A/Residues: 2845/3; 2863/3; 2887/3; 2899/3; 2922/1; 2985/1; 3008/1; 3065/1
A/Keywords: alternative splicing; cell binding; coiled coil; connective tissue; disulfide
F/1-23/Domains: signal sequence #status predicted <SIG>
F/24-3124/Product: collagen alpha 1(XI) chain #status predicted <MAT>
F/24,1189-3124/Product: collagen alpha 1(XII) chain short splice form #status predicted
F/24-114/Domains: IIIA #status predicted <IIIA>
F/24-105/Domains: fibronectin type III repeat homology <FN3A>
F/137-301/Domains: von Willebrand factor type A repeat homology <VWA1>
F/333-425/Domains: IIIB #status predicted <IIIB>
F/333-414/Domains: fibronectin type III repeat homology <FN3B>
F/437-601/Domains: von Willebrand factor type A repeat homology <VWA2>
F/623-1178/Domains: IIIC #status predicted <IIIC>
F/630-711/Domains: fibronectin type III repeat homology <FN3C>
F/7721-802/Domains: fibronectin type III repeat homology <FN3D>
F/812-895/Domains: fibronectin type III repeat homology <FN3E>
F/905-986/Domains: fibronectin type III repeat homology <FN3F>
F/1095-1076/Domains: fibronectin type III repeat homology <FN3G>
F/11086-1169/Domains: fibronectin type III repeat homology <FN3H>
F/1137-1361/Domains: von Willebrand factor type A repeat homology <VWA3>
F/11384-12295/Domains: IIID #status predicted <IIID>
F/11384-1465/Domains: fibronectin type III repeat homology <FN3I>
F/1474-1557/Domains: fibronectin type III repeat homology <FN3J>
F/1566-1647/Domains: fibronectin type III repeat homology <FN3K>
F/1565-1738/Domains: fibronectin type III repeat homology <FN3L>
F/1756-1838/Domains: fibronectin type III repeat homology <FN3M>
F/1847-1928/Domains: fibronectin type III repeat homology <FN3N>
F/1937-2019/Domains: fibronectin type III repeat homology <FN3O>
F/2028-2110/Domains: fibronectin type III repeat homology <FN3P>
F/2119-2139/Domains: fibronectin type III repeat homology <FN3Q>
F/2207-2294/Domains: fibronectin type III repeat homology <FN3R>
F/2335-2490/Domains: von Willebrand factor type A repeat homology <VWA4>
F/2438-2440/Domains: cell adhesion #status predicted
F/2509-2150/Domains: IX, homologous to NC4 domain of type IX collagen #status predicted
F/2751-2902/Domains: collagenous CO2 #status predicted <COL2>
F/2899-2901/Domains: cell attachment (R-G-D) motif
F/2903-2945/Domains: non-collagenous NC2 #status predicted <NC2>
F/2946-3048/Domains: non-collagenous CO1 #status predicted <COL1>
F/3049-3124/Domains: non-collagenous NC1 #status predicted <NC1>
F/3121,1006,1032,1094,1512,1767,2210,2273,2532,2683/Binding site: carbohydrate (Asn)
F/3280,2789,2836,2842,2860,2866,2869,3004,3007/Modified site: hydroxyproline (Pro) #statu

```
Query Match      13.9%; Score 125; DB 2; Length 3124;  
Best Local Similarity 27.1%; Pred. No. 0.028;  
Matches    48; Conservative   39; Mismatches   74; Indels   16; Gaps   9;
```

OY

1 DLYPLDSSGV-LHNMNEIYYFVEQLAHKF--ISPO-PRMSFYVSRGRTLLKLT--E 54
| : | : | | : : : | : : : : : : : : : : : :

```

D5 2327 DIVELTQASMSIGDGNPFRKVVKFEVFNTVGAFADLLINPAGIIGVSLQVYSLEMAOSEKFLMTFD 2386
O5 55 DREGIROGLEBLOKAVLPGQDPTVHMEGEPASQIY-YENRQGYRTASVYIYLTGCELIHED 113
D5 2387 DKQAQALALQNVQ--YNGNTRTSKALTFIIEKVLJTBESGRRGRVCPKLVVVTVDGRSGDE 2444
O5 114 LFFSSEKRNRSRLGALIVCVGKVDNENQALAIAD--SKDHFEPVNDGQALOGI 168
D5 2445 V----RKCAIVYIQSGSFVFGVADVDVYNELAKIASKPSRRHAFVDD- FDAQEKI 2486

```

RESULT 4

immunodominant microneme protein Etpl00 - *Eimeria tenella*

C:\Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

R;Tomley, F.M.; Clarke, L.B.; Kawazoe, U.; Dijkema, R.; Kok, J.J.

A;Title: Sequence of the gene encoding an immunodominant

A;Acceßion: A45638

A: Molecule type: DNA

A:Cross-reference: UNIPROT:043981: GB:AF032905: NID:Q2707732: PTDN:AAD03350
M/KEBIDEB: 1-12 <10M>

A;Note: Sequence extracted from NCBI database (NCBIN: //132, NCBI: //136)
E:48-218/Domain: von Willebrand factor type A repeat homology <VWA1>

E;238-296/Domain: thrombospondin type 1 repeat homolog <THRL>
E;308-371/Domain: thrombospondin type 1 repeat homolog <THR3>

EF372-432/Domain: thrombospondin type 1 repeat homology <THR3>

F;494-556/Domain: thrombospondin type 1 repeat homology <THR5:

1. The first group of respondents (n = 10) was composed of individuals who had been employed by the company for less than one year. This group was selected to represent new employees who were likely to have limited experience with the company's policies and procedures.

Best Local Similarity 25.6%; Pred. No. 0.0097;

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

27

20 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 10

103 I A I D G B L H D L F I S E R E A N K S K U L G A I V I C V G - N D F N E I Q I A K I

DB 158 LVMIDGA--SNLPSQIKSSAALRDAGALVVLGVSGVNSSECKSL

SECRET

RWHUIC

N;Alternate names: leukocyte adhesion receptor p150,95 alpha

C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change

R;Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.

A;Reference number: A36584

A;Accession: A36584

A;Residues: 1-1163 <COR>

A;Note: this revision to the sequence from reference A35543 in

J. Biol. Chem. 265, 2782-2788, 1990

A:Reference number: A35543: MUTD:90153906: PMID:2303426

A;ACCESSION: A35543

A/Molecule type: DNA
 A/Residues: 1-834 <CO2>
 A/Note: this sequence has been revised in reference A36584
 R/Corbi, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.
 EMBL J. 6, 4033-4028, 1987
 A/Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocyte
 A/Reference number: S00864; MUID:88166645; PMID:3327687
 A/Accession: S00864
 A/Molecule type: mRNA
 A/Residues: 1-755; 'L', 757-1163 <CO3>
 A/Cross-references: GB:M1695; EMBL:Y00093; NID:G487829; PIND:AAA59180.1; PID:G487830
 A/Note: part of this sequence was confirmed by protein sequencing
 C/Comment: A common beta chain (CD18) forms a heterodimer with CD1b to form Mac-1 on my
 C/Genetics:
 A/Gene: GDB:ITGAX; CD11C
 A/Cross-references: GDB:119758; OMIM:151510
 A/Map position: 16p11.2-16p11.2
 C/Superfamily: cell surface glycoprotein CD1b; von Willebrand factor type A repeat hom
 C/Keywords: calcium; cell adhesion; glycoprotein; heterodimer; magnesium; tandem repeat;
 F:1-19/Domains: signal sequence #status predicted <SIG>
 F:20-1163/Product: cell surface glycoprotein CD1c #status predicted <MAT>
 F:20-1107/Domains: extracellular #status predicted <EXT>
 F:149-319/Domains: von Willebrand factor type A repeat homology <VWA4>
 F:1108-1133/Domains: transmembrane #status predicted <TM>
 F:1134-1163/Domains: intracellular #status predicted <INT>
 F:61,89,392,697,735,899,939,1050/Binding site: carbohydrate (Asn) (covalent) #status pre

Query Match 13.5%; Score 121.5; DB 1; Length 1163;
 Best Local Similarity 25.1%; Pred. No. 0.017;
 Matches 47; Conservative 35; Mismatches 74; Indels 31; Gaps 9;

QY 1 DIVFLIDKSGSVLHHNNEIYFVEQLAHKISPOLMSFVSTRTGTLTKLTDDE-- 57
 D 151 DIVFLIDKSGSVLHHNNEIYFVEQLAHKISPOLMSFVSTRTGTLTKLTDDE-- 210
 QY 58 ---QIRGGLBELQVLPFGSDTYMEGPERASEQIYENRQGYRAS-VIATLTDGELHED 113
 D 211 NPLSLASVHQLQ-----GFTYTATLQNVNHRFLPHASGARRATKILIVITDGKEGD 265
 QY 114 LFFYSERANRSDLAIVYCVGK-----KDFNETQLARIAD--SKDHVFPVNDG 161
 D 266 SLDYKD-VIPMDAGIIRYALIGVLAFOGNSWKEIND-----LASKPQEHIFKVED- 318
 QY 162 FOALQGI 168
 D 319 FDALKDI 325

RESULT 6

S00551

leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse
 N/Alternate names: complement-3 receptor alpha chain
 C/Species: Mus musculus (house mouse)
 C/Date: 30-Sep-1989 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
 C/Accession: S00551; 159078

R/Pytela, R.
 EMBL J. 7, 1371-1378, 1988
 A/Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the 1
 A/Reference number: S00551; MUID:88312584; PMID:3044779
 A/Accession: S00551
 A/Molecule type: DNA
 A/Residues: 1-1153 <PYT>

A/Cross-references: UNIPROT:P05555; EMBL:X07640; NID:G52982; PIND:CAA30479.1; PID:G52983
 A/Note: the authors translated the codon CAC for residue 569 as Gln
 R/Sastre, L.; Roman, J.M.; Teplov, D.B.; Dreyer, W.J.; Gee, C.E.; Larson, R.S.; Roberts,
 Proc. Natl. Acad. Sci. U.S.A. 83, 5644-5648, 1986
 A/Title: A partial genomic DNA clone for the alpha subunit of the mouse complement recep
 A/Reference number: 159078; MUID:86287312; PMID:2942940
 A/Accession: 159078
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 11-44 <RES>
 A/Cross-references: GB:M14293; NID:g198993; PIND:AAA3484.1; PID:g554193

C/Genetics:
 A/Gene: Mac-1
 C/Superfamily: cell surface glycoprotein CD1b; von Willebrand factor type A repeat homo
 C/Keywords: cell adhesion; glycoprotein; transmembrane protein
 F:1-16/Domains: signal sequence #status predicted <SIG>
 F:17-1153/Product: leukocyte surface glycoprotein Mac-1 alpha chain #status experimental
 F:148-318/Domains: von Willebrand factor type A repeat homology <VWA2>
 F:1106-1129/Domains: transmembrane #status predicted <TM>

Query Match 13.4%; Score 120.5; DB 2; Length 1153;
 Best Local Similarity 27.2%; Pred. No. 0.021;
 Matches 52; Conservative 34; Mismatches 66; Indels 39; Gaps 12;

QY 1 DIVFLIDKSGSVLHHNNEIYFVEQLAHKIS---POLMSFVSTRTGTLTKLTD-- 55
 D 150 DIVFLIDKSGSVLHHNNEIYFVEQLAHKIS---POLMSFVSTRTGTLTKLTD-- 197
 QY 56 ---REQIRGGLBELQVLP-----GGDTYMEGPERASEQIYENRQGYRASVIATL 106
 D 198 IHTFTNDFKKNPSPRSHVSIKOLNGRTKTASGIRKVRKELFKTNGARENAAKILVIT 257
 QY 107 DGEIHEDLFFYSE--REANRSDLAIVYCVGKD--FNETQLARIAD-----SKDHVFP 157
 D 258 DGEKFGDPLDYKQVPEADRA---GVIRYIVGVGNFKNRQSRRELDITASKPAGEHVFO 314
 QY 158 VNDGFOALQGI 168
 D 315 V-DNFEALNTI 324

RESULT 7

C2M5

classical-complement-pathway C3/C5 convertase (EC 3.4.21.43) C2 component precursor - mor
 N/Alternate names: C3 convertase; C5 convertase; complement C2
 C/Species: Mus musculus (house mouse)
 C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
 C/Accession: A38876; B36593; I54429

R/Ishikawa, N.; Nonaka, M.; Wetzel, R.A.; Colten, H.R.
 submitted to Genbank, January 1991
 A/Reference number: A38875
 A/Accession: A38876
 A/Molecule type: DNA
 A/Residues: 1-760 <IS2>

A/Cross-references: UNIPROT:P21180; GB:M57891; GB:J05661; NID:g192436; PIND:AAA63294.1; I
 R/Ishikawa, N.; Nonaka, M.; Wetzel, R.A.; Colten, H.R.
 J. Biol. Chem. 265, 19040-19046, 1990
 A/Title: Murine complement C2 and factor B genomic and cDNA cloning reveals different me
 A/Reference number: A36593; MUID:91035430; PMID:2229060
 A/Accession: B36593

A/Molecule type: mRNA
 A/Residues: 1-760 <ISH>
 A/Cross-references: EMBL:M57891; NID:g192436; PIND:AAA63294.1; PID:g192437
 R/Falut, A.; Wakeland, E.K.; McConnell, T.J.; Gitlin, J.; Whitehead, A.S.; Colten, H.R.
 Immunogenetics 25, 290-298, 1987

A/Title: DNA polymorphism of MHC III genes in inbred and wild mouse strains.
 A/Reference number: I54429; MUID:87192938; PMID:2883115
 A/Accession: I54429
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 660-677; 'R', 679,681-723; 'G', 725 <RES>
 A/Cross-references: GB:M16271; NID:g199289; PIND:AAA39562.1; PID:g199290

C/Genetics:
 A/Intons: 16/1; 91/1; 153/1; 212/1; 245/3; 290/3; 337/1; 384/1; 414/1; 461/1; 492/3; 53;
 C/Complex: The proenzyme forms a complex with C4a and is activated by cleavage into C2a,
 C/Function:
 A/Description: cleaves complement C3 and complement C5 alpha chains

A/Superfamily: complement B/C2; complement factor H repeat homology; tryptain homology; v
 C/Keywords: alternative splicing; complement classical pathway; duplication; glycoprotein
 F:1-18/Domains: signal sequence #status predicted <SIG>
 F:19-250/Product: complement C2b fragment #status predicted <CB>
 F:22-89/Domains: complement factor H repeat homology <FH1>
 F:94-149/Domains: complement factor H repeat homology <FH2>

R/Christiano, A.M.; Greenpan, D.S.; Lee, S.; Uitto, J.
 J. Biol. Chem. 269, 20256-20262, 1994
 A/Title: Cloning of human type VII collagen. Complete primary sequence of the alpha (VII)
 A/Reference number: A54849; MUID:94327588; PMID:8051117
 A/Accession: A54849
 A/Status: not compared with conceptual translation
 A/Molecule type: mRNA
 A/Residues: 1-2944 <CHR>
 A/Cross-references: UNIPROT:Q02388; GB:L02870; NID:9987124; PIDN:AAA75438.1; PID:9987125
 R/Tanaka, T.; Takahashi, K.; Furukawa, F.; Imamura, S.
 Biochem. Biophys. Res. Commun. 183, 958-963, 1992
 A/Title: Molecular cloning and characterization of type VII collagen cDNA.
 A/Reference number: F00844; MUID:92231902; PMID:15671409
 A/Accession: F00844
 A/Molecule type: mRNA
 A/Residues: 340-475; 'A', 'L', 'S', 'T', 'A', 'S', 'H', 'S', 'T', 'L', 'C', 'M', 'B', 'A', 'T', 'R', 'W', 'H', 'P', 'C', 'N', 'R', 'G', 'S', 'H', 'T', 'R', 'A', 'C', 'E', 'P', 'N', 'R', 'P', 'A', 'R', 'A', 'G', '524-528, 'C', 'A', 'C', 'R', 'O', 'S', 'S', 'R', 'E', 'F', 'E', 'R', 'E', 'N', 'C', 'E', 'S', 'D', 'D', 'B', 'J', 'D', 'I', '1', '5', '2', 'D', 'B', 'J', 'D', 'I', '3', '6', '9', '4', 'N', 'I', 'D', '9', '4', '5', '3', '6', '9', '8', 'P', 'I', 'D', 'N', 'B', 'A', 'A', '2', '8', '5', '1', 'P', 'I', 'D', '9', '4', '5', '3', '6', '9', '9'
 A/Experimental source: keratinocyte
 A/Note: the authors translated the codon ACC for residues 394 and 397 as Tyr
 R/Parente, M.G.; Chung, L.C.; Rymaszewski, J.; Woodley, D.T.; Wynn, K.C.; Bauer, E.A.; Mac
 proc. Natl. Acad. Sci. U.S.A. 88, 6931-6935, 1991
 A/Title: Human type VII collagen: cDNA cloning and chromosomal mapping of the gene.
 A/Reference number: S16316; MUID:91334380; PMID:18711109
 A/Accession: S16316
 A/Molecule type: mRNA
 A/Residues: 815-892, 'E', '894-1439 <PAR>
 A/Cross-references: GB:M65158; GB:S949017; NID:9180914; PIDN:AAA96439.1; PID:9180915
 A/Experimental source: keratinocyte
 R/Gemmon, W.R.; Abernethy, M.L.; Peddila, K.M.; Pristavnik, P.S.; Cook, M.E.; Wright, J.
 J. Invest. Dermatol. 99, 691-696, 1992
 A/Title: Noncollagenous (NC1) domain of collagen VII resembles multidomain adhesion prot
 A/Reference number: I56328; MUID:91107742; PMID:1469284
 A/Accession: I56328
 A/Status: translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 372-517, 'D', 'V', '520-540, 'W', '542-1255 <RES>
 A/Cross-references: GB:S51236; NID:9262308; PIDN:AA824637.1; PID:9262309
 R/Selkater, J.L.; Eisen, A.Z.; Bauer, E.A.; Morris, N.P.; Glanville, R.W.; Burgeson, R.E.
 J. Biol. Chem. 264, 3822-3826, 1989
 A/Title: Cleavage of type VII collagen by interstitial collagenase and type IV collagenase
 A/Reference number: A30296; MUID:89139437; PMID:2537292
 A/Accession: A30296
 A/Molecule type: protein
 A/Residues: 'A', '1240-1246, 'G', '1248-1250, 'X', '1253-1255, 'Q', '1257, 'E', '2032, 'C', '2034-2041; '
 A/Note: two reported peptides cannot be reliably located
 R/Greenpan, D.S.
 Hum. Mol. Genet. 2, 273-278, 1993
 A/Title: The carboxyl-terminal half of type VII collagen, including the non-collagenous
 A/Reference number: I48103; MUID:93271985; PMID:8499916
 A/Accession: I48103
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 2395-2871, 'G', '2873-2944 <RES>
 A/Cross-references: GB:L06862; NID:9388713; PIDN:AAA89196.1; PID:9388714
 R/Christiano, A.M.; Rymaszewski, M.; Uitto, J.
 Proc. Natl. Acad. Sci. U.S.A. 91, 3549-3553, 1994
 A/Title: Dominant dystrophic epidermolysis bullosa: identification of a Gly --> Ser sub
 A/Reference number: A55255; MUID:94224777; PMID:8170945
 A/Accession: A55255
 A/Contents: annotation
 C/Comment: prolines and lysines at the third position of the tripeptide repeating unit
 ed and subsequently O-glycosylated.
 C/Genetic:
 A/Genes: GDB:COL7A1; EBR1; EBD1; EB
 A/Cross-references: GDB:128750; OMIM:120120
 A/Map position: 3p21.3-3p21.3
 A/Note: defects in this gene can result in dominant and recessive dystrophic epidermolys
 A/Note: there are 118 introns
 C/Complex: type VII collagen is probably a homotrimer
 C/Function:
 A/Description: structural component of extracellular polymer associated with anchoring f
 C/Keywords: coiled coil; extracellular matrix; glycoprotein; hydroxyllysine; hydroxyproli
 F/1-16/Domain: signal sequence #status predicted <SIG>
 F/17-2944/Product: collagen alpha 1(VII) chain #status predicted <MAT>

F/17-1253/Domain: amino-terminal nonhelical #status predicted <NC1>
 F/36-201/Domain: von Willebrand factor type A repeat homology <VWA1>
 F/231-318/Domain: fibronectin type III repeat homology <FN1>
 F/327-413/Domain: fibronectin type III repeat homology <FN2>
 F/414-502/Domain: fibronectin type III repeat homology <FN3>
 F/508-593/Domain: fibronectin type III repeat homology <FN4>
 F/598-683/Domain: fibronectin type III repeat homology <FN5>
 F/686-771/Domain: fibronectin type III repeat homology <FN6>
 F/776-862/Domain: fibronectin type III repeat homology <FN7>
 F/864-952/Domain: fibronectin type III repeat homology <FN8>
 F/954-1045/Domain: fibronectin type III repeat homology <FN9>
 F/1052-1219/Domain: von Willebrand factor type A repeat homology <VWA2>
 F/1170-1172/Region: cell attachment (R-G-D) motif
 F/1189-1253/Region: cysteine/proline-rich
 F/1254-2783/Region: interrupted helical
 F/1334-1336/Region: cell attachment (R-G-D) motif
 F/2008-2010/Region: cell attachment (R-G-D) motif
 F/2553-2555/Region: cell attachment (R-G-D) motif
 F/2784-2944/Domain: carboxyl-terminal nonhelical #status predicted <NC2>
 F/2876-2929/Domain: animal Kunitz-type proteinase inhibitor homology <BPI>
 F/3137-786, 1109/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F/3167-2176, 2185, 2188, 2664, 2667, 2673/Modified site: 4-hydroxyproline (Pro) #
 F/2625-2631/Modified site: 5-hydroxylysine (Lys) #status experimental
 F/2625-2631/Binding site: carbohydrate (Lys) (covalent) #status experimental
 F/2634, 2804/Disulfide bonds: interchain #status predicted

Query Match 11.3%; Score 102; DB 2; Length 2944;
 Best Local Similarity 26.8%; Pred. No. 2.6;
 Matches 48; Conservative 34; Mismatches 79; Indels 18; Gaps 10;

Qy 1 DLYFIDKSGSV-LHWNELIYFVEOLAHKF--ISFO-LMSFIVSTRGTT--LMKL 52
 Db 38 DIVFLDSSSIGRNSRFRVSLFLEGLVLPFGSAASQGVAFVAVQVSDPRTFGLDAL 97

Qy 53 TDRSQRLQGLERLQKTLPGDDTYMHGFEASQIYYENRQGRISVIALTDGLH 112
 Db 98 GSGGDVIR-AIREIS-YKGNRTGTAALLHVAADHFLPOLARDGVPCVCLITDGR-SQ 153

Qy 113 DLFFYSREANRSHDLGAIYCVGQFNETOLARIND--SKDHFVPNDGFOALGII 169
 Db 154 DLV--DTMAORLKGQGVKLPAVGIRKADPELRVASQPSDFPFVND-FSLIRILL 208

RESULT 14
 BBHU
 Complement factor B precursor [validated] - human
 N/Alternate names: C3 convertase; C3 proactivator; glycine-rich beta-glycoprotein; heat-
 N/Contents: alternative-complement-pathway C3/C5 convertase (EC 3.4.21.47) Bb Fragment
 C/Species: Homo sapiens (man)
 C/Date: 19-Feb-1984 #sequence revision 05-Aug-1994 #text change 09-Jul-2004
 C/Accession: S34075; A44622; R00934; A19188; A19947; B25971; S14339; A44628; 154
 R/Mojta, J.E.; Jahn, I.; de la Salle, H.; Hauptmann, G.
 Submitted to the EMBL Data Library, March 1993
 A/Reference number: S34075
 A/Accession: S34075
 A/Molecule type: mRNA
 A/Residues: 1-764 <ME3>
 A/Cross-references: UNIPROT:P00751; EMBL:X72875; NID:9297568; PIDN:CAA51389.1; PID:92975.
 R/Woods, D.B.; Markham, A.F.; Ricker, A.T.; Goldberg, G.; Colten, H.R.
 Proc. Natl. Acad. Sci. U.S.A. 79, 5661-5665, 1982
 A/Title: Isolation of cDNA clones for the human complement protein factor B, a class III
 A/Reference number: A44622; MUID:83039428; PMID:6957884
 A/Accession: A44622
 A/Molecule type: mRNA
 A/Residues: 467-546; 550-595; 752-764 <MOO>
 A/Cross-references: GB:J00185; GB:J00186
 A/Note: the authors translated the codon TAC at 519 as Thr; the nucleic acid translation
 R/Mole, J.E.; Anderson, J.K.; Davison, E.A.; Woods, D.E.
 J. Biol. Chem. 259, 3407-3412, 1984
 A/Title: Complete primary structure for the zymogen of human complement factor B.
 A/Reference number: A20751; MUID:84161997; PMID:6546754
 A/Accession: A20751
 A/Molecule type: protein; mRNA

A/Residues: 26-764 <MOL>
 A/Cross-references: GB:K01566
 A/Note: nucleic acid translation differs from the sequence shown in having 300-Leu, 328-
 A/Note: 736-Ser was also found
 A/Note: glycosylation sites were determined
 R/Character: D.L.; Gagnon, J
 Biochem. J. 209, 61-70, 1983
 A/Title: Amino acid sequence of the Bb fragment from complement factor B. Sequence of th
 A/Reference number: A19188; MUID:83204002; PMID:6342610
 A/Content: the final paper in a series documenting the sequence, glycosylation site, an
 A/Accession: A19188
 A/Molecule type: protein
 A/Residues: 260-236, 'T', 298-764 <CHR>
 R/Campbell, R.D.; Porter, R.R.
 Proc. Natl. Acad. Sci. U.S.A. 80, 4464-4468, 1983
 A/Title: Molecular cloning and characterization of the gene coding for human complement
 A/Reference number: A19947; MUID:83272641; PMID:6308626
 A/Accession: A19947
 A/Molecule type: DNA
 A/Residues: 346-764 <CAM>
 A/Cross-references: GB:U00125
 A/Accession: B19947
 A/Molecule type: mRNA
 A/Residues: 339-509 <CA1>
 A/Cross-references: GB:U00126; NID:g187723; PIDN:AAA36226.1; PID:g553536
 R/Wu, L.; Morley, B.J.; Campbell, R.D.
 Cell 48, 331-342, 1987
 A/Title: Cell-specific expression of the human complement protein factor B gene: evidenc
 A/Reference number: A25971; MUID:87102880; PMID:3643061
 A/Accession: B25971
 A/Molecule type: DNA
 A/Residues: 1-99 <MUL>
 A/Cross-references: GB:M15082; NID:g187699; PIDN:AAA59625.1; PID:g553534
 R/Niemann, M.A.; Bhown, A.S.; Miller, E.J.
 Biochem. J. 274, 473-480, 1991
 A/Title: The principal site of glycation of human complement Factor B.
 A/Reference number: 814339; MUID:91174758; PMID:2006911
 A/Accession: 814339
 A/Molecule type: protein
 A/Residues: 270-329 <NIE>
 A/Note: binding site for carbohydrate to lysine under artificial conditions
 R/Morley, B.J.; Campbell, R.D.
 EMBO J. 3, 153-157, 1984
 A/Title: Internal homologues of the Ba fragment from human complement component factor B
 A/Reference number: A44628; MUID:84158524; PMID:6323161
 A/Accession: A44628
 A/Molecule type: preliminary
 A/Status: preliminary
 A/Residues: 16-225, 'F', 227-259 <MOR>
 R/Schwedde, W.; Lutrig, B.; Sokolowski, T.; Ebelier, C.; Weiss, E.H.; Meyer zum Busche
 Immunobiology 188, 221-232, 1993
 A/Title: Human complement factor B: functional properties of a recombinant zymogen of th
 A/Reference number: 154409; MUID:94041399; PMID:8225386
 A/Accession: 154409
 A/Status: translated from GB/EMBL/DBD
 A/Molecule type: mRNA
 A/Residues: 1-764 <RES>
 A/Cross-references: GB:867310; NID:g452937; PIDN:AA13989.1; PID:g4261689
 R/Horlrich, T.; Kim, S.; Matsunoto, M.; Watanabe, I.; Fujita, S.; Volanakis, J.E.
 Mol. Immunol. 30, 1587-1592, 1993
 A/Title: Human complement factor B: cDNA cloning, nucleotide sequencing, phenotypic con
 A/Reference number: 157824; MUID:94067177; PMID:8247029
 A/Accession: 157824
 A/Status: translated from GB/EMBL/DBD
 A/Molecule type: mRNA
 A/Residues: 1-31, 'O', 33-764 <RE2>
 A/Cross-references: GB:U15702; NID:g291921; PIDN:AAA16820.1; PID:g291922
 C/Comment: 292-Cys has a free sulphydryl.
 C/Genetics:
 A/Status: GDB:BF
 A/Cross-references: GDB:119726; OMIM:138470
 A/Map position: 6p21.3-6p21.3
 A/Intons: 21/3; 99/3; 346/1; 390/1; 424/1; 470/1; 502/3; 542/1; 593/2; 619/1; 652/3; 65

A/Note: the list of introns may be incomplete
 A/Note: gene is located in the major histocompatibility complex, class III region
 A/Complex: complement factor B initially forms an inactive complex with complement factor
 ment factor C3b forming active C3/C5 convertase; Ba is released
 C/Function:
 A/Description: Bb is a serine protease; C3/C5 convertase cleaves complement C3 alpha ct
 A/Pathway: complement alternate pathway
 C/Superfamily: complement B/C2; complement factor H repeat homology; trypsin homology; v
 C/Keywords: acute phase; complement alternate pathway; duplication; glycoprotein; hydroly
 F:1-25/Domain: signal sequence #status predicted <SIG>
 F:26-764/Product: complement factor B #status experimental <MAT>
 F:26-259/Product: complement factor Ba fragment #status experimental <BAF>
 F:37-98/Domain: complement factor H repeat homology <FH1>
 F:103-158/Domain: complement factor H repeat homology <FH2>
 F:165-218/Domain: complement factor H repeat homology <FH3>
 F:260-764/Product: C3/C5 convertase Bb fragment #status experimental <BBF>
 F:268-458/Domain: von Willebrand factor type A repeat homology <VFA>
 F:482-752/Domain: trypsin homology #status atypical <TRY>
 F:37-76, 62-98, 103-145, 131-158, 165-205, 191-218, 478-596, 511-527, 599-615, 656-682, 695-725/Dic
 F:122, 142, 285, 378/Binding site: carbohydrate (asn) (covalent) #status experimental
 F:259-260/Cleavage site: Arg-Lys (complement factor D) #status experimental
 F:526, 576, 699/Active site: His, Asp, Ser #status experimental

Query Match 11.3%; Score 101.5; DB 1; Length 764;
 Best Local Similarity 22.5%; Pred. No. 0.59;
 Matches 47; Conservative 43; Mismatches 74; Indels 45; Gaps 10;

QY 1 DLYFIIDKSGSV-----LHHMNEIYFVBOIAKFIISPOLMSFIYSTRGTTLMKTE- 54
 DB 270 NITVLVDGSDISASFTGAKCLVNIIEKVASYGKPE-RYGLVYATIPKIVAKVSEA 327
 QY 55 ---DREQIRQGEEL-----QKVLPGDITVYHGEFERNASEQIYENR-----QGY-RTA 99
 DB 328 DSNNAWVTFKQLNEIYVBDHKLSGNT-----KALQAVYSWMSWPDVPEGNVTR 381
 QY 100 SYVIALTDG-----ELHEDLFYSREANRSDLCALYICVG--VKDFNETQ 144
 DB 382 HVIILMTDGLHMGDPITVIDEIRDLITGKGRKPRBEDYLDVYVGVPVNVQVINA 441
 QY 145 LARIADSKDHPVNDGFOALGGIHSIL 173
 DB 442 LASKDNEQHFVKVD-MENTLEDVYQMI 469

RESULT 15
 T28797
 hypothetical protein C16B9.1 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 29-Oct-1999 #sequence_revlsion 29-Oct-1999 #text_change 09-Jul-2004
 C/Accession: T28797
 R/Geisels, C.
 submitted to the EMBL Data Library, November 1995
 A/Description: The sequence of C. elegans cosmid C16B9.
 A/Reference number: Z20525
 A/Accession: T28797
 A/Status: preliminary; translated from GB/EMBL/DBD
 A/Molecule type: DNA
 A/Residues: 1-567 <GB1>
 A/Cross-references: UNIPROT:Q18048; EMBL:U99677; PIDN:AA047957.1; GSPDB:GN00028; CESP:CI
 A/Experimental source: strain Bristol N2; clone C16B9
 C/Genetics:
 A/Status: CESP:C16B9.1
 A/Map position: X
 A/Intons: 15/3; 77/3; 132/1; 185/3; 231/1; 271/2; 301/2; 430/3; 509/1

Query Match 11.3%; Score 100.5; DB 2; Length 567;
 Best Local Similarity 24.1%; Pred. No. 0.5;
 Matches 41; Conservative 38; Mismatches 72; Indels 19; Gaps 9;

QY 1 DLYFIIDKSGSVLHHMNEIYFVBOIAKFIISPOLMSFIYSTRGTTLMKTE 56
 DB 393 DIGIITDSSGSLEKNFQKQLAVBOQPIISDNRIVGIVPAGKTKRVLANFSQNK 452

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 19:37:47 ; Search time 58.0753 Seconds
(without alignment)
1525.429 Million cell updates/sec

Title: US-09-970-076-2_COPY_44_216
Perfect score: 899
Sequence: 1 DLYFIDKSGSVLHMNEIY.....HFPVNDGFGALQGIHSIL 173

Scoring table: BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 segs, 512079187 residues
Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03.*
1: uniProt_sprot.*
2: uniProt_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	899	100.0	564	1 ATRI_HUMAN	Q96X22 homo sapien
2	890	99.0	562	1 ATRI_MOUSE	Q9CZ52 mus musculu
3	519	57.7	489	1 ATRI_HUMAN	P58335 homo sapien
4	517	57.5	487	2 O6DFX2	O6DFX2 mus musculu
5	348	38.7	641	2 O8BVW2	O8BVW2 mus musculu
6	141	15.7	1332	2 O9BPQ8	Q9BPQ8 halocynthia
7	138.5	15.4	1161	1 ITAD_RAT	Q9GYE7 rattus norv
8	129.5	14.4	724	2 O04588	O04588 elmeria max
9	128.5	14.3	1152	1 ITAM_HUMAN	P11215 homo sapien
10	128.5	14.3	3119	1 CAIC_MOUSE	O60847 mus musculu
11	125	13.9	441	2 O8T6T5	O8T6T5 mytilus edu
12	125	13.9	444	2 O8T5C3	O8T5C3 mytilus edu
13	125	13.9	3124	1 CAIC_CHICK	P13944 gallus galli
14	124	13.8	765	2 O9U8J9	Q9U8J9 neospora ca
15	122.5	13.6	1162	1 ITAD_HUMAN	O13349 homo sapien
16	122	13.6	453	2 O8T5C2	O8T5C2 mytilus gal
17	121.5	13.5	712	2 O43981	O43981 elmeria ten
18	121.5	13.5	1163	1 ITAX_HUMAN	P20702 homo sapien
19	120.5	13.4	1153	1 ITAM_MOUSE	P05555 mus musculu
20	118.5	13.2	3063	1 CAIC_HUMAN	Q99715 homo sapien
21	117	13.0	1182	2 O8C6K9	O8C6K9 mus musculu
22	117	13.0	1823	2 O7PRP5	O7PRP5 anopheles g
23	116.5	13.0	760	2 O70350	O70350 mus musculu
24	116.5	13.0	920	2 Q28984	Q28984 sus scrofa
25	116	12.9	637	2 O81VX1	O81VX1 homo sapien
26	116	12.9	1284	2 O6P159	O6P159 homo sapien
27	116	12.9	1329	1 KP10_HUMAN	Q99218 homo sapien
28	113.5	12.6	517	2 O43853	O43853 homo sapien
29	113.5	12.6	760	1 CO2_MOUSE	P21180 mus musculu
30	112	12.5	588	2 O6ZED8	O6ZED8 synecococc
31	111.5	12.4	2104	2 Q21281	Q21281 caenorhabdi

32	111.5	12.4	2104	2	Q96AN4	Q96AN4 caenorhabdi
33	111	12.3	440	2	Q8CBT2	Q8CBT2 mus musculu
34	111	12.3	848	2	Q8C720	Q8C720 mus musculu
35	111	12.3	3567	2	Q9ES77	Q9ES77 mus musculu
36	110	12.2	1626	2	O8NFW1	O8NFW1 homo sapien
37	110	12.2	1888	1	CAIC_CHICK	P12018 gallus galli
38	109	12.1	3183	2	O65ZC2	O65ZC2 caenorhabdi
39	109	12.1	3767	1	MU43_CAEEL	P34576 caenorhabdi
40	108	12.0	1188	1	ITAX_MOUSE	P61622 mus musculu
41	108	12.0	1188	2	O7TQC3	O7TQC3 mus musculu
42	107.5	12.0	599	2	O8MVQ1	O8MVQ1 bolitena vi
43	107	11.9	919	2	O7SR52	O7SR52 lymnaea ata
44	107	11.9	2944	2	O63870	O63870 mus musculu
45	106.5	11.8	415	2	Q8DGY6	Q8DGY6 synecococc

ALIGNMENTS

RESULT 1
ID ATRI_HUMAN STANDARD; PRT; 564 AA.
AC Q96X22; Q96P02; Q9NVP3;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Anthrax toxin receptor 1 precursor (Tumor endothelial marker 8).
GN Name=ANTXR1; Synonyms=ATR, TEM8;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCB1_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=20407466; PubMed=10947988; DOI=10.1126/science.289.5482.1197;
RA St Croix B., Rago C., Velculescu V., Traverso G., Romans K.E.,
RA Montgomery E., Lal A., Riggs G.J., Lengauer C., Vogelstein B.,
RA Kinzler K.W.;
RT "Genes expressed in human tumor endothelium";
RL Science 289:1197-1202(2000).
[2]
RP SEQUENCE FROM N.A. (ISOFORM 2), AND INTERACTION WITH ANTHRAX TOXIN.
RX MEDLINE=21557240; PubMed=11700562; DOI=10.1038/n510198;
RA Bradley K.A., Mogridge J., Mourez M., Collier R.J., Young J.A.T.;
RT "Identification of the cellular receptor for anthrax toxin.";
RL Nature 414:225-229(2001).
[3]
RP SEQUENCE OF 184-564 FROM N.A. (ISOFORM 1), AND SEQUENCE FROM N.A.
(ISOFORM 3).
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakematsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohtsushi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Nagahashi K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Komoto H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA Omura Y., Abe K., Kamihara K., Katsuma N., Sato K., Tanikawa M.,
RA Yamazaki M., Nishimura K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hirakawa S., Chiba Y.,
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yoshida M., Houcha T.,
RA Kuwano Y., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Togawa S., Komai F., Hara R., Takeuchi K., Arita M.,
RA Iinose N., Musashino K., Yuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshihawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Motiya S., Nomiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiro M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,

RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mitushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs";
RL Nat. Genet. 36:40-45(2004).
RN [4]
RN SEQUENCE FROM N.A. (ISOFORM 4).
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Strophton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Ueda T.B., Toshlyak S., Carrinci P., Prange C.,
RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RN INTERACTION WITH ANTHRAX TOXIN.
RC TISSUE=Placenta;
RX MEDLINE=22608610; PubMed=12700348; DOI=10.1073/pnas.0431098100;
RA Scobie H.M., Rainey G.J.A., Bradley K.A., Young J.A.T.;
RT "Human capillary morphogenesis protein 2 functions as an anthrax toxin
RT receptor";
RL Proc. Natl. Acad. Sci. U.S.A. 100:5170-5174(2003).
RN [6]
RN SPICE ISOFORM(S) THAT ARE POTENTIAL NMD TARGET(S).
RX PubMed=14759258; DOI=10.1186/gb-2004-5-2-r8;
RA Hallman R.T., Green R.E., Brenner S.E.;
RT "An unappreciated role for RNA surveillance";
RL Genome Biol. 5:RESEARCH008.1-RESEARCH008.16(2004).
CC -1- FUNCTION: Cellular role is not yet known.
CC -1- SUBUNIT: Binds to the protective antigen (PA) of *Bacillus*
CC anthracis. Binding does not occur in the presence of calcium.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Probable).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Comment=Experimental confirmation may be lacking for some
CC isoforms;
CC Name=1;
CC IsoId=Q9H6X2-1; Sequence=Displayed;
CC Note=May be produced at very low levels due to a premature stop
CC codon in the mRNA, leading to nonsense-mediated mRNA decay;
CC Name=2;
CC IsoId=Q9H6X2-2; Sequence=VSP_000444, VSP_000445;
CC Name=3;
CC IsoId=Q9H6X2-3; Sequence=VSP_000446, VSP_000447;
CC Name=4;
CC IsoId=Q9H6X2-4; Sequence=VSP_000448, VSP_000449;
CC -1- TISSUE SPECIFICITY: Highly expressed in tumor endothelial cells
CC but not in normal endothelial cells.
CC -1- DOMAIN: Binding to PA seems to be effected through the VMA domain.
CC -1- SIMILARITY: Belongs to the ATR family.
CC -1- SIMILARITY: Contains 1 VMA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>/
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL: AF279145; AAK52094.1; -
DR EMBL: AF421380; AAL26496.1; -
DR EMBL: AK025429; BAB15128.1; ALT_INIT.
DR EMBL: AK001463; BAA91707.1; ALT_FRAME.
DR EMBL: BC012074; AAI12074.1; -
DR Genew: HGNC:21014; ANTXRL.
DR H-InvDB: HIX000125; -
DR MIM: 606410; -
DR InterPro: IPR008400; Anth_Ig.
DR InterPro: IPR008399; Ant_C.
DR InterPro: IPR02035; VWF_A.
DR Pfam: PF05587; Anth_Ig_1.
DR Pfam: PF05566; Ant_C_1.
DR Pfam: PF00092; VMA_1.
DR SMART: SM00327; VMA_1.
DR PROSITE: PS50234; VWF_A.
KV Alternative splicing; Glycoprotein; Receptor; Signal; Transmembrane.
FT SIGNAL 1 32
FT CHAIN 1 32
FT DOMAIN 33 321
FT TRANSMEM 322 342
FT DOMAIN 343 564
FT DOMAIN 44 215
FT DOMAIN 360 368
FT DOMAIN 506 564
FT CARBOHYD 166 166
FT CARBOHYD 184 184
FT CARBOHYD 262 262
FT VARSPLIC 365 368
FT VARSPLIC 369 564
FT VARSPLIC 268 297
FT VARSPLIC 298 564
FT VARSPLIC 319 333
FT VARSPLIC 334 564
SQ SEQUENCE 564 AA; 62789 MW; B118A00D5DF2233 CRC64;
Query Match 100.0%; Score 899; DB 1; Length 564;
Best Local Similarity 100.0%; Pred. No. 16-69;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DLVFIIDKSGSVLHHNNEIYFVEQLAHKFIISQLMSFVFSRGTLMKLTEDRQIR 60
DB 44 DLVFIIDKSGSVLHHNNEIYFVEQLAHKFIISQLMSFVFSRGTLMKLTEDRQIR 103
QY 61 QGLEBELQKLVPGGDTVMHSEGFERASQIYYENQGRYASVIALTDGELHEDLFFYSER 120
DB 104 QGLEBELQKLVPGGDTVMHSEGFERASQIYYENQGRYASVIALTDGELHEDLFFYSER 163
QY 121 EANRSRDLAGIYVCVGVKDFNETQLARIADSKDHFVPVNDGFQALQGIHSIL 173
DB 164 EANRSRDLAGIYVCVGVKDFNETQLARIADSKDHFVPVNDGFQALQGIHSIL 216
RESULT 2
ATRI_MOUSE STANDARD; PRT; 562 AA.
AC Q9C252;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)

RT "Human capillary morphogenesis protein 2 functions as an anthrax toxin receptor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:5170-5174(2003).
 RN [3]
 RP SEQUENCE OF 78-489 FROM N.A. (ISOFORM 3), AND SEQUENCE FROM N.A. (ISOFORM 4).
 RP TISSUE=Synovial cell;
 RC PubMed:14702039; DOI=10.1038/ng1285;
 RX Ota T., Suzuki Y., Nishikawa T., Ohtsuki T., Sugiyama T., Irie R., Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H., Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S., Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y., Shiratori K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikawa E., Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M., Nimomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori Y., Tanai H., Kimata M., Watanabe S., Yoshida M., Horita T., Iihida S., Ono Y., Takiguchi S., Watanabe S., Yoshida M., Horita T., Kusanagi J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O., Nomura Y., Togaiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N., Musashino K., Yuuki H., Ooshima A., Sasaki N., Aotsuka S., Yoshihara Y., Matsunawa H., Ichihara T., Shiohara T., Sano S., Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O., Moriya S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H., Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B., Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y., Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiyara T., Oho T., Yamada K., Fujii Y., Ozaki K., Hiro M., Ohmori Y., Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S., Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsunura K., Nakajima Y., Mizuno T., Morinaga M., Komatsu T., Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T., Mizushima-Shigino J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K., Okumura K., Nagase T., Nomura N., Kikuchi H., Maeno Y., Yamashita R., Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
 RT "Complete sequencing and characterization of 21,243 full-length human cDNAs.";
 RL Nat. Genet. 36:40-45(2004).
 CC -1- FUNCTION: Cellular role is not yet known.
 CC -1- SUBUNIT: Binds to the protective antigen (PA) of Bacillus anthracis in a divalent cation-dependent manner, with the following preference: calcium > manganese > magnesium > zinc.
 CC Seems to bind to collagen type IV and laminin.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (isoforms 1 and 2). Secreted (isoform 3). Isoform 1 is expressed at the cell surface while isoform 2 is predominantly expressed within the endoplasmic reticulum and not at the plasma membrane.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=4;
 CC Name=1;
 CC IsoId=P58335-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=P58335-2; Sequence=VSP_008343;
 CC Name=3;
 CC IsoId=P58335-3; Sequence=VSP_008344, VSP_008345;
 CC Name=4;
 CC Note=No experimental confirmation available;
 CC Name=4;
 CC IsoId=P58335-4; Sequence=VSP_008346;
 CC Note=No experimental confirmation available;
 CC -1- TISSUE SPECIFICITY: Expressed in colon, heart, kidney, lung, liver, peripheral blood leukocytes, placenta, skeletal muscle, small intestine and spleen.
 CC -1- DOMAIN: Binding to PA seems to be effected through the VWA domain.
 CC -1- SIMILARITY: Belongs to the ATR family.
 CC -1- SIMILARITY: Contains 1 VMPA domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>)

CC or send an email to license@ebi.ac.uk.
 CC -----
 DR EMBL: AY040326; AAT7722.1; -;
 DR EMBL: AY233452; AAP04016.1; -;
 DR EMBL: AK055636; BAB70976.1; ALT_INT.
 DR EMBL: AK091721; BAC03731.1; -;
 DR GenBank: HGNC:21732; ANTXR2.
 DR MIM: 608041; -;
 DR InterPro: IPR008399; Ant_C.
 DR InterPro: IPR002035; VWF_A.
 DR Pfam: PF05586; Ant_C; 1.
 DR Pfam: PF00092; VWF_1.
 DR PROSITE: PSS0234; VWF_1.
 KM Alternative splicing; Glycoprotein; Receptor; Signal; Transmembrane.
 FT SIGNAL 1 33
 FT CHAIN 34 489
 FT DOMAIN 34 318
 FT TRANSMEM 319 341
 FT DOMAIN 342 489
 FT DOMAIN 44 213
 FT CARBOHYD 250 250
 FT CARBOHYD 260 260
 FT VARSPLIC 213 315
 FT VARSPLIC 290 322
 FT VARSPLIC 323 489
 FT VARSPLIC 477 489
 FT CONFLICT 357 357
 FT SEQUENCE 489 AA; 53692 MW; B9F679DB75B6E2B7 CRC64;
 Query Match 57.7%; Score 519; DB 1; Length 489;
 Best Local Similarity 59.5%; Pred. No. 9.7e-37;
 Matches 103; Conservative 34; Mismatches 34; Indels 2; Gaps 1;
 QY 1 DLVFIIDKSGSVVHHNNIEYFVEQLAHKRPISQILNPFVSTRTGLMKLTEDREQIR 60
 DB 44 DLVFIIDKSGSVVHHNNIEYFVEQLAHKRPISQILNPFVSTRTGLMKLTEDREQIR 103
 QY 61 QGLLEIQLVLPQGDITVHGEFASRISQIYYENRQYRTASVIALTDGELHEDLPFYSER 120
 DB 104 KGLIEDLRVSPVGETYIHBEGLKLANQI--QKAGKLTSSIIIALTDGKLDGLVPSYAEK 161
 QY 121 EAKRSRDLGAIYVCVGVKDPNEQLARIADSKSHVPPVNDGFQALGIIHSIL 173
 DB 162 EAKISRSLSGASVCGVLDPEQALRIADSKSHVPPVNDGFQALGIIHSIL 214
 RESULT 4
 Q6DFX2 ID Q6DFX2 PRELIMINARY; PRT; 487 AA.
 AC Q6DFX2; 25-OCT-2004 (T-EMBLrel. 28, Created)
 DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)
 DE Anthrax toxin receptor 2.
 GN Name=Antxr2;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Eye; DOI=10.1073/pnas.242603899;
 RX MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold B.A., Grouse L.H., Dege J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altshuler S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatzenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.J., Scheetz T.E.,
 RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loguigliano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Boock S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Holyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fehey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywicki M.I., Skelton U., Smallos D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CS7BL/6; TISSUE=Eye;
 RA Strauberg R.;
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC076595; AAH76595.1; -
 DR GO: GO:0016021; C:integral to membrane; IEA.
 DR GO: GO:0004872; F:receptor activity; IEA.
 DR InterPro: IPR008400; Antch_IG.
 DR InterPro: IPR008399; Ant_C_IG.
 DR InterPro: IPR02035; VWF_A.
 DR Pfam: PF05587; Antch_IG; I.
 DR Pfam: PF05586; Ant_C_IG; I.
 DR Pfam: PF00092; VMA; I.
 DR SMART: SM00327; VMA; I.
 DR PROSITE: PS50234; VWF_A; I.
 KM Receptor.
 SQ SEQUENCE 487 AA; 53184 MW; 61A400D60BC8DE59 CRC64;
 Query Match 57.5%; Score 517; DB 2; Length 487;
 Best Local Similarity 59.5%; Pred. No. 1,46-36;
 Matches 103; Conservative 31; Mismatches 37; Indels 2; Gaps 1;
 QY 1 DLYPIIDKSGSVLHNNHNIYFVEQLAHKFIISPOLBMSFIVFSTRTTKLTEDRQIR 60
 DB 44 DLYPIIDKSGSVLHNNHNIYFVEQLAHKFIISPOLBMSFIVFSTRTTKLTEDRQIR 103
 QY 61 OGLEBLQKLVPGCDTYVHGEFPERASBOIYYENRQGYRTASYIILTDGELHEDLPFYSER 120
 DB 104 KGLEBLQKLVPGCDTYVHGEFPERASBOIYYENRQGYRTASYIILTDGELHEDLPFYSER 161
 QY 121 EANRSRDGLAIYVCVGVKDFNETQLARIADSKCHVPVNDGFPALQGIHSITL 173
 DB 162 EAKKSRELGLASVYCVGVLDPEQQLERLADSKCHVPVNDGFPALQGIHSITL 214
 RESULT 5
 ID O8BVM2 PRELIMINARY; PRT; 641 AA.
 AC O8BVM2;
 DT 01-MAR-2003 (TEMBLrel. 23, Created)
 DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
 DE Mus musculus adult male testis cDNA, RIKEN full-length enriched
 DE library, clone:933430J1 product:hypothetical Prolin-rich region/von
 DE Willebrand factor type A domain containing protein, full insert
 DE sequence.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CS7BL/6J; TISSUE=Testis;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;

RT "High-efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CS7BL/6J; TISSUE=Testis;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CS7BL/6J; TISSUE=Testis;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CS7BL/6J; TISSUE=Testis;
 RX MEDLINE=20493374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.,
 RT "Normalization and subcloning of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CS7BL/6J; TISSUE=Testis;
 RX MEDLINE=20530913; PubMed=11078861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Kono H., Akiyama J., Ishii K., Kiteunai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kaibuchi K.,
 RA Fujisake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,
 RA Okazaki Y., Ishikawa T., Ozawa K., Tanaka T., Matsubara S., Kawai J.,
 RA Yonezaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CS7BL/6J; TISSUE=Testis;
 RA Adachi J., Aizawa K., Akimura T., Arai K., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirose T.,
 RA Hori F., Imocani K., Ishii Y., Itoh M., Kagawa I., Kasuawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohnishi N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akaiura S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK077206; BAC36683.1; -
 DR GO: GO:0016021; C:integral to membrane; IEA.
 DR GO: GO:0004872; F:receptor activity; IEA.
 DR InterPro: IPR008400; Antch_IG.
 DR InterPro: IPR002035; VWF_A.
 DR Pfam: PF05587; Antch_IG; I.
 DR Pfam: PF00092; VMA; I.
 DR PRINTS: PR00453; VWFADOMAIN.
 DR SMART: SM00327; VMA; I.
 DR PROSITE: PS50234; VWF_A; I.
 KM Hypothetical protein.
 SQ SEQUENCE 641 AA; 70415 MW; 1993300730B85E3 CRC64;
 Query Match 38.7%; Score 348; DB 2; Length 641;
 Best Local Similarity 43.8%; Pred. No. 9,8e-22;
 Matches 74; Conservative 30; Mismatches 65; Indels 0; Gaps 0;
 QY 1 DLYPIIDKSGSVLHNNHNIYFVEQLAHKFIISPOLBMSFIVFSTRTTKLTEDRQIR 60

```

Db      76 DLYVLDKSSGVADNMWHTIHSFASGLVKKFTNPNLRISITITYSREAVIILLPLTSDSKEIN 135
Qy      61 QGLEELKVLPGGDTYMEGFERASSEOIYYENROGYRTASVIALTDGELHEDLPFYSER 120
Db      136 KSLIVLKSIYQGLTHMQKLRKANEGJRKSTLGGRIYNSVIALTDLGLLLKRYLPTME 195
Qy      121 EARSRLGAIYVCVGVKDFRETOLARIADSKHVPFPPNPGFOLQGIT 169
Db      196 EAKKARHMGALIVYGVFMYSKQOLVNIAGDPDRCFGVDEGFSALLEGV 244

RESULT 6
Q9BP08 ID Q9BP08 PRELIMINARY; PRT; 1332 AA.
AC Q9BP08 01-JUN-2001 (TREMBLERel. 17, Created)
DT 01-JUN-2001 (TREMBLERel. 17, Last sequence update)
DE Integrin alpha Hrl precursor.
GN Name=HrTGA1;
OS Haemocytus roretzi (Sea squirt).
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Scolidobranchia; Pyuridae; Haemocytus.
OX NCBI_TaxID=7729;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hemocyte;
RC MEDLINE=21103187; PubMed=11160215;
RA Miyazawa S., Azumi K., Nonaka M.;
RT "Cloning and characterization of integrin alpha subunits from the
RT solitary ascidian, Haemocytus roretzi."
RL J. Immunol. 166:1710-1715(2001).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
DR EMBL: AB048261; BAB21479.1; -.
DR HSSP: P11215; 1BHO.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0008305; C:integrin complex; IEA.
DR GO: GO:0005515; F:protein binding; IEA.
DR GO: GO:0007160; F:cell-matrix adhesion; IEA.
DR GO: GO:0007229; P:integrin-mediated signaling pathway; IEA.
DR InterPro: IPR000413; Integrin_alpha.
DR InterPro: IPR002035; VWF_A.
DR Pfam: PF00092; VMA; 1.
DR PRINTS: PRO1185; INTEGRINA.
DR SMART: SM00453; VWFADOMAIN.
DR SMART: SM00327; VMA; 1.
DR PROSITE: PS50234; VWF_A; 1.
DR Cell adhesion; Integrin; Signal; Transmembrane.
FT SIGNAL 1 1332 potential.
FT CHAIN 31 1332 Integrin alpha Hrl.
SQ SEQUENCE 1332 AA; 145851 MW; 0D9108D2805CFEAE CRC64;

Query Match 15.7%; Score 141; DB 2; Length 1332;
Best Local Similarity 24.3%; Pred. No. 0.0023;
Matches 51; Conservative 34; Mismatches 59; Indels 66; Gaps 7;

Qy      1 DLYVLDKSSGVADNMWHTIHSFASGLVKKFTNPNLRISITITYSREAVIILLPLTSDSKEIN 60
Db      206 DLYVLDKSSGVADNMWHTIHSFASGLVKKFTNPNLRISITITYSREAVIILLPLTSDSKEIN 243
Qy      61 QGLEELKVLPGGDTYMEGFERASSEOIYYENROGYRT-----AKDLIGKEIVR 98
Db      244 VGVVQYSHYVEGKINKOKITTEISIGFPLDLPFNADVRI---QLQGYTTYGRALQ 300
Qy      99 -----ASVIALTDGELHEDLPFYSERANRSRDGAIYVCVGVKDFRETOL 145
Db      301 KVIKRPDPAIYGNKQVILLTLDGAKDNKILP--NANRLNKGIATFAVGVEYDISEL 358
Qy      146 ARIA---DSKDHVFPVNDGFOLQGITHSI 172

```

```

Db      359 KLIASGDSIDRYFTYTD--FGELDSIVKSL 387

RESULT 7
ID ITAD RAT STANDARD; PRT; 1161 AA.
AC Q9OYE7;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DE Integrin alpha-D precursor.
GN Name=Itgad;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA O'Brien M.M., VanderVlieten M., Kilgannon P.D., Dietrich G.,
RA Gallatin W.M.;
RT "Cloning of rat alpha D, a novel beta 2 integrin."
RL Submitted (JUG-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Integrin alpha-D/beta-2 is a receptor for ICAM3 and
CC VCAM1. May play a role in the atherosclerotic process such as
CC clearing lipoproteins from plaques and in phagocytosis of blood-
CC borne pathogens, particulate matter, and senescent erythrocytes
CC from the blood (By similarity).
CC -1- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-D
CC associates with beta-2 (By similarity).
CC -1- DOMAIN: The Integrin I-domain (Inert) is a VWF_A domain. Integrins
CC with I-domains do not undergo protease cleavage.
CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
CC -1- SIMILARITY: Contains 1 VWF_A domain.

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF021334; AAF21241.1; -.
DR HSSP: P11215; 1BHO.
DR InterPro: IPR000413; Integrin_alpha.
DR InterPro: IPR002035; VWF_A.
DR Pfam: PF01839; FG-GAP; 3.
DR Pfam: PF00357; Integrin_alpha; 1.
DR Pfam: PF00092; VMA; 1.
DR PRINTS: PRO1185; INTEGRINA.
DR PRINTS: PRO0453; VWFADOMAIN.
DR SMART: SM00191; Int_alpha; 5.
DR SMART: SM00327; VMA; 1.
DR PROSITE: PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE: PS50234; VWF_A; 1.
DR Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
KM Repeat; signal; Transmembrane.
FT SIGNAL 1 19 potential.
FT CHAIN 20 1161 Extracellular (Potential).
FT DOMAIN 1101 1121 Potential.
FT DOMAIN 1122 1161 Cytoplasmic (Potential).
FT REPEAT 34 87 FG-GAP 1.
FT REPEAT 86 88 FG-GAP 2.
FT DOMAIN 152 334 VWF_A.
FT REPEAT 352 402 FG-GAP 3.
FT REPEAT 403 454 FG-GAP 4.
FT REPEAT 456 517 FG-GAP 5.
FT REPEAT 519 577 FG-GAP 6.
FT REPEAT 582 634 FG-GAP 7.

```


FT CA_BIND 467 475 Potential.
 FT CA_BIND 531 539 Potential.
 FT CA_BIND 594 602 Potential.
 FT SITE 1126 1130 GPCR motif.
 FT DISULFID 69 76 By similarity.
 FT DISULFID 108 126 By similarity.
 FT DISULFID 656 711 By similarity.
 FT DISULFID 769 775 By similarity.
 FT DISULFID 845 860 By similarity.
 FT DISULFID 993 1017 By similarity.
 FT DISULFID 1022 1027 By similarity.
 FT CARBOHYD 61 61 N-linked (GlcNAc...)
 FT CARBOHYD 89 89 N-linked (GlcNAc...)
 FT CARBOHYD 245 245 N-linked (GlcNAc...)
 FT CARBOHYD 393 393 N-linked (GlcNAc...)
 FT CARBOHYD 696 696 N-linked (GlcNAc...)
 FT CARBOHYD 734 734 N-linked (GlcNAc...)
 FT CARBOHYD 784 784 N-linked (GlcNAc...)
 FT CARBOHYD 907 907 N-linked (GlcNAc...)
 FT CARBOHYD 936 936 N-linked (GlcNAc...)
 FT CARBOHYD 1045 1045 N-linked (GlcNAc...)
 SQ SEQUENCE 1161 AA; 12600 MW; 2258491A984A705E CRC64;

Query Match 15.4%; Score 138.5; DB 1; Length 1161;
 Best Local Similarity 28.3%; Pred. No. 0.0032;
 Matches 52; Conservative 36; Mismatches 71; Indels 25; Gaps 10;

QY 1 DLVYILKSGSV-LHNNIEIYFVEQLAHKFIQPLMNSFVFSRGTLMKLE----- 54
 DB 152 DIAFLIGSSINORDPAQMDPFAVMGEPASTSTLSLMQSYNLIKHTFTFEPKNIL 211
 QY 55 DREQIRGELBELQVLPDGTVMHEGFERASEQIYYENRQGYRTA-SVTLALTDGELHED 113
 DB 212 DPQSLVDPYVQLQ-----GLTYTATGIRTYMBELFHNKNGSRKAKKILVITDGGQKRT 266
 QY 114 LFFYSE--REANSRDIGAIVCVGV--FNE-TQLARI-----ADSKDHVPVNDGFOA 164
 DB 267 PLRYSVDYIPADKA---GIIRYAIQVGDAPQEPALXELNTIGSAPQDHVFKVGN-FAA 322
 QY 165 LOGI 168
 DB 323 LRSL 326

RESULT 8
 Q04588 PRELIMINARY; PRT; 724 AA.

AC Q04588; TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Major antigen homologous sequence (empi00).
 OS Eimeria maxima.
 OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae;
 OC Eimeria.
 NCBI_TaxID=5804;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=3119203; PubMed=8426611; DOI=10.1016/0166-6851(93)90255-V;
 RA Paramontes L.E., Hug D., Huembelin M., Weber G.;
 RT "Sequence of a major Eimeria maxima antigen homologous to the Eimeria
 tenella microneme protein Etp100.";
 RT Mol. Biochem. Parasitol. 57:171-174(1993).
 DR EMBL; M99058; AAA28076.1; -.
 DR PIR; A48569; A48569.
 DR HSP; P07996; ILSL.
 DR GO; GO:0008083; F: growth factor activity; IEA.
 DR InterPro; IPR000762; PTN MK.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF00090; TSP_1/ 5.
 DR Pfam; PF00092; VMA; 1.
 DR PRINTS; PR00453; VWFADOMAIN.

DR SMART; SM00209; TSP1; 6.
 DR SMART; SM00327; VMA; 1.
 DR PROSITE; PS50092; TSP1; 5.
 DR PROSITE; PS50234; VMA; 1.
 SQ SEQUENCE 724 AA; 75808 MW; AC2A0E7A346A7B3E CRC64;

Query Match 14.4%; Score 129.5; DB 2; Length 724;
 Best Local Similarity 28.1%; Pred. No. 0.011;
 Matches 48; Conservative 30; Mismatches 68; Indels 25; Gaps 10;

QY 1 DLVYILKSGSV-LHNNIEIYFVEQLAHKFIQPLMNSFVFSRGTLMKLE----- 57
 DB 47 DVLIVVESGSIQSNIGKRSFISNPAQTPSLSPDVRGLVYFGSAVTRMDLSRA 106
 QY 58 QIRGELBELQVLP--GDTVMHEGFERASEQIYYENRQGYR--TASVTLALTDGELHED 113
 DB 107 QMDLAAAKKLPYAGSTYTHGLAKA-ELIFSPQKGRDPAKPKILVMTDGA----- 161
 QY 114 LFFYSEB-----ANSRDIGAIVCVGV--KDFNETQLARIA--DSKDHV 155
 DB 162 -----SSRSQTLISAERKLRNRGVIIIVLVGVGTGVSASACRSIAGCDTSDTV 208

RESULT 9

ITAM HUMAN
 ID ITAM HUMAN STANDARD; PRT; 1152 AA.

AC P11215;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Integrin alpha-M precursor (CD11b) surface glycoprotein MAC-1 alpha
 DE subunit (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1)
 DE (Neutrophil adherence receptor).
 GN Name: ITGAM; Synonyms: CD11B, CR3A;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88315033; PubMed=2457584;
 RA Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;
 RT "The human leukocyte adhesion glycoprotein Mac-1 (complement receptor
 RT type 3, CD11b) alpha subunit. Cloning, primary structure, and relation
 RT to the integrins, von Willebrand factor and factor B.";
 RT J. Biol. Chem. 263:12403-12411(1988).
 RL [2]
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88190151; PubMed=2833753;
 RA Arnout M.A., Remold-O'Donnell B., Pierce M.W., Harris P., Tenen D.G.;
 RT "Molecular cloning of the alpha subunit of human and guinea pig
 RT leukocyte adhesion glycoprotein MO1: chromosomal localization and
 RT homology to the alpha subunits of Integrins.";
 RT Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).
 RL [3]
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88257215; PubMed=2454931; DOI=10.1083/jcb.106.6.2153;
 RA Arnout M.A., Gupta S.K., Pierce M.W., Tenen D.G.;
 RT "Amino acid sequence of the alpha subunit of human leukocyte adhesion
 RT receptor MO1 (complement receptor type 3).";
 RT J. Cell Biol. 106:2153-2158(1988).
 RL [4]
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93123746; PubMed=8419480;
 RA Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.F., Tenen D.G.;
 RT "Structural analysis of the CD11b gene and phylogenetic analysis of
 RT the alpha-integrin gene family demonstrate remarkable conservation of
 RT genomic organization and suggest early diversification during
 RT evolution.";
 RT J. Immunol. 150:480-490(1993).
 RL [5]
 RN [5]
 RP SEQUENCE OF 9-1153 FROM N.A.
 RX MEDLINE=89098893; PubMed=2563162;

RA Hickey D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L.,
RA Roth G.J.;
RT "CDNA sequence for the alpha M subunit of the human neutrophil
RT adherence receptor indicates homology to integrin alpha subunits";
RL Proc. Natl. Acad. Sci. U.S.A. 86:257-261(1989).
RN [6]
RP SEQUENCE OF 1-9 FROM N.A.
RX MEDLINE=92073318; PubMed=1683702;
RA Shelley C.S., Arnaut M.A.;
RT "The promoter of the CD1b gene directs myeloid-specific and
RT developmentally regulated expression.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:10525-10529(1991).
RN [7]
RP SEQUENCE OF 1-9 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=92144986; PubMed=1346576;
RA Paul H.L., Rosmarin A.G., Tenan D.G.;
RT "Characterization of the myeloid-specific CD1b promoter.";
RL Blood 79:865-870(1992).
RN [8]
RP SEQUENCE OF 17-31.
RX MEDLINE=87076671; PubMed=3539202; DOI=10.1016/0167-4838(86)90037-3;
RA Pierce M.W., Remold-O'Donnell E., Todd R.F. III, Arnaut M.A.;
RT "N-terminal sequence of human leukocyte glycoprotein Mol: conservation
RT across species and homology to platelet IIb/IIIa.";
RL Biochim. Biophys. Acta 874:368-371(1986).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 148-331.
RX MEDLINE=9517458; PubMed=7867070; DOI=10.1016/0092-8674(95)90517-0;
RA Lee J.O., Rieu P., Arnaut M.A., Liddington R.C.;
RT "Crystal structure of the A domain from the alpha subunit of integrin
RT CR3 (CD11b/CD18).";
RL Cell 80:631-638(1995).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 148-334.
RX MEDLINE=96363671; PubMed=8747460; DOI=10.1016/S0969-2126(01)00271-4;
RA Lee J.O., Barker L.A., Arnaut M.A., Liddington R.C.;
RT "Two conformations of the integrin A-domain (1-domain): a pathway for
RT activation?";
RL Structure 3:1333-1340(1995).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 148-337.
RX MEDLINE=98362596; PubMed=9687375; DOI=10.1016/S0969-2126(98)00093-8;
RA Baldwin E.T., Server R.W., Bryant G.L., Jr., Curry K.A.,
RA Fairbanks M.B., Finzel B.C., Garlick R.L., Heinrichson R.L.,
RA Horton N.C., Kelley L.L., Mildner A.M., Moon J.B., Mott J.E.,
RA Matzner V.T., Tomich C.S., Watenpaugh K.D., Wiley V.H.;
RT "cation binding to the integrin CD11b I domain and activation model
RT assessment.";
RL Structure 6:923-935(1998).
RN [12]
RP 3D-STRUCTURE MODELING OF 17-616.
RX MEDLINE=98226734; PubMed=9560195; DOI=10.1073/pnas.95.9.4870;
RA Ouyig C., Springer T.A.;
RT "Experimental support for a beta-propeller domain in integrin alpha-
RT subunits and a calcium binding site on its lower surface.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875(1998).
RN [13]
RP FUNCTION: Integrin alpha-M/beta-2 is implicated in various
RP adhesive interactions of monocytes, macrophages and granulocytes
RP as well as in mediating the uptake of complement-coated particles.
RP It is identical with CR-3, the receptor for the IC3b fragment of
RP the third complement component. It probably recognizes the R-G-D
RP peptide in C3b. Integrin alpha-M/beta-2 is also a receptor for
RP fibrinogen, factor X and ICAM1. It recognizes P1 and P2 peptides
RP of fibrinogen gamma chain.
RN [14]
RP SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-M
RP associates with beta-2.
RN [15]
RP SUBCELLULAR LOCATION: Type I membrane protein.
RN [16]
RP TISSUE SPECIFICITY: Predominantly expressed in monocytes and
RP granulocytes.
RN [17]
RP DOMAIN: The integrin I-domain (insert) is a VWA domain. Integrins
RP with I-domains do not undergo protease cleavage.
RN [18]
RP SIMILARITY: Belongs to the integrin alpha chain family.

CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
CC -1- SIMILARITY: Contains 1 VWA domain.
CC DATABASE: NAME=PROW; NOTE=CD guide CD11b entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11b.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, J03925; AAAS9544.1; -;
DR EMBL, M18044; AAAS9491.1; -;
DR EMBL, J04145; AAAS9903.1; -;
DR EMBL, S52227; AAB24821.1; -;
DR EMBL, S52152; AAB24821.1; JOINED.
DR EMBL, S52153; AAB24821.1; JOINED.
DR EMBL, S52154; AAB24821.1; JOINED.
DR EMBL, S52155; AAB24821.1; JOINED.
DR EMBL, S52157; AAB24821.1; JOINED.
DR EMBL, S52159; AAB24821.1; JOINED.
DR EMBL, S52161; AAB24821.1; JOINED.
DR EMBL, S52164; AAB24821.1; JOINED.
DR EMBL, S52165; AAB24821.1; JOINED.
DR EMBL, S52167; AAB24821.1; JOINED.
DR EMBL, S52169; AAB24821.1; JOINED.
DR EMBL, S52170; AAB24821.1; JOINED.
DR EMBL, S52173; AAB24821.1; JOINED.
DR EMBL, S52174; AAB24821.1; JOINED.
DR EMBL, S52180; AAB24821.1; JOINED.
DR EMBL, S52181; AAB24821.1; JOINED.
DR EMBL, S52184; AAB24821.1; JOINED.
DR EMBL, S52189; AAB24821.1; JOINED.
DR EMBL, S52191; AAB24821.1; JOINED.
DR EMBL, S52203; AAB24821.1; JOINED.
DR EMBL, S52192; AAB24821.1; JOINED.
DR EMBL, S52212; AAB24821.1; JOINED.
DR EMBL, S52213; AAB24821.1; JOINED.
DR EMBL, S52216; AAB24821.1; JOINED.
DR EMBL, S52219; AAB24821.1; JOINED.
DR EMBL, S52220; AAB24821.1; JOINED.
DR EMBL, S52221; AAB24821.1; JOINED.
DR EMBL, S52222; AAB24821.1; JOINED.
DR EMBL, S52225; AAB24821.1; JOINED.
DR EMBL, M76724; AAAS6410.1; -;
DR EMBL, M64477; AAAS1960.1; -;
DR PIR, A31108; RWH01B.
DR PDB, 1A8X; Model; @17-1152.
DR PDB, 1BHO; X-ray; 1/2=-.
DR PDB, 1BHO; X-ray; 1/2=-.
DR PDB, 1IDN; X-ray; 1/2=-.
DR PDB, 1IDJ; X-ray; @140-331.
DR PDB, 1JLM; X-ray; @143-334.
DR PDB, 1MLU; X-ray; A=137-331.
DR PDB, 1MF7; X-ray; A=144-337.
DR PDB, 1MF7; X-ray; A=140-335.
DR PDB, 1MA5; X-ray; A=144-345.
DR Genew; HGNC:6149; ITGM.
DR MIM, 120980; -;
DR GO, GO:0008305; C:integrin complex; TAS.
DR GO, GO:0007155; P:cell adhesion; TAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01859; FG-GAP_3.
DR Pfam; PF00357; Integrin_alpha; 1.
DR Pfam; PF00092; VWA; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFDOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.

DR PROSITE; PSS0234; VMPA. 1.
 KW 3D-structure; Calcium; Cell adhesion; Direct protein sequencing;
 KW Transmembrane. 1
 FT SIGNAL 1 16
 FT CHAIN 17 1152 Integrin alpha-M.
 Query Match 14.3%; Score 128.5; DB 1; Length 1152;
 Best Local Similarity 26.7%; Pred. No. 0.024;
 Matches 51; Conservative 38; Mismatches 63; Indels 39; Gaps 11;
 QY 1 DLVFIIDKSGSVL-HHNNELIYFVEQLAHKFIISPOLMSFTVSTRTTLMKTED---- 55
 DB 150 DIAFLIDGSGSIIIPDRFRKKEFVST-----VMEQLKSKTLFS-----LMQYSEFRHH 199
 QY 56 -----REQIRQGLSELQKVLPGCDYMHGCFERASBOIYYENRGYRTA-SVITALT 106
 DB 200 FTFRKPPNNNPISLVKPIQOL--GRTHATGIRKVRRELFNTNARKAKAPILIVIT 257
 QY 107 DGEIHEDLPFYSE--REANSRDGAIVCGVKNFNETOLAR-----IADS--KDHVP 157
 DB 258 DGEKFGDPLGVEVDVPEADRE---GVIRYVIGVGDARFSEKSRQELMTIASKPPRDHVFQ 314
 QY 158 VINDFQALOGI 168
 DB 315 VNN-PEALKTI 324
 RESULT 10
 ID CALC_MOUSE STANDARD; PRT; 3119 AA.
 AC 060847; P70322;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Collagen alpha 1(XII) chain precursor.
 GN Name=Col12a1;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING (ISOFORMS XIIA-1 AND XIIA-2).
 RC STRAIN=C57BL/6J, and Swiss Webster; TISSUE=Skin;
 RX MEDLINE=96170761; PubMed=8601036;
 RA Boehme K., Li Y., Oh P.S., Olsen B.R.;
 RT "Primary structure of the long and short splice variants of mouse collagen XII and their tissue-specific expression during embryonic development.";
 RL Dev. Dyn. 204:432-445(1995).
 RN [2]
 RP PARTIAL SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING (ISOFORMS XIIA-2 AND XIIA-2).
 RC STRAIN=C57BL/6J; TISSUE=Skin fibroblast;
 RX MEDLINE=99348349; PubMed=10419532; DOI=10.1074/jbc.274.31.22053;
 RA Kania A.M., Reichenberger E., Baur S.T., Karimux N.Y., Taylor R.W., Olsen B.R., Nishimura I.;
 RT "Structural variation of type XII collagen at its carboxyl-terminal NCI domain generated by tissue-specific alternative splicing.";
 RL J. Biol. Chem. 274:22053-22059(1999).
 CC -1- FUNCTION: Type XII collagen interacts with type I collagen-containing fibrils, the COL1 domain could be associated with the surface of the fibrils, and the COL2 and NC3 domains may be localized in the pericellular matrix (By similarity).
 CC -1- SUBUNIT: Trimer of identical chains each containing 190 kDa of nontriple-helical sequences (By similarity).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=4;
 CC Comment=The final tissue form of collagen XII may contain homotrimers or any combination of the various isoforms;
 CC Name=XIIA-1;
 CC IsoId=Q60847-1; Sequence=Displayed;

CC Name=XIIA-2; Synonym=ERK;
 CC IsoId=Q60847-2; Sequence=VSP_001151, VSP_001152;
 CC Name=XIIB-1;
 CC IsoId=Q60847-3; Sequence=VSP_001150;
 CC Name=XIIB-2;
 CC IsoId=Q60847-4; Sequence=VSP_001150, VSP_001151, VSP_001152;
 CC -1- TISSUE SPECIFICITY: Highest expression in tendons, perichondrium, skin, cornea, sclera, blood vessels, and peritoneum.
 CC -1- DEVELOPMENTAL STAGE: The long NC3 XIIA isoforms are predominant at early stages (ED7 and 11); at later stages of development (ED15 and 17) the short NC3 XIIB forms become the major forms. As the short NC3 forms become the major product, the long splice variant continues to be expressed in several tissues, even after birth.
 CC The long NC1 isoforms, XIIA-1 and XIIB-1, peak in 15-day old embryos and decrease in 17-day old ones. The expression of the short NC1 form XIIB-2 remains constant throughout late stages of embryonic development (ED15 and ED17).
 CC -1- PTM: The triple-helical tail is stabilized by disulfide bonds at each end (By similarity).
 CC -1- PTM: Prolines at the third position of the tripeptide repeating unit (G-X-Y) are hydroxylated in some or all of the chains (By similarity).
 CC -1- PTM: O-glycosylation of isoform XIIA-2; glycosaminoglycan of chondroitin-sulfate type (By similarity).
 CC -1- SIMILARITY: Belongs to the fibril-associated collagens with interrupted helices (FACIT) family.
 CC -1- SIMILARITY: Contains 18 fibronectin type III domains.
 CC -1- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
 CC -1- SIMILARITY: Contains 4 VMPA domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U25652; AAA99719.1; ALT_SEQ.
 DR EMBL; U57095; AAB07047.1; --
 DR HSSP; P18614; LMHP.
 DR WGD; MG1:88448; Col12a1.
 DR Interpro; IPR008160; Collagen.
 DR Interpro; IPR008985; Cona_1like_1ec_g1.
 DR Interpro; IPR003961; FN_III.
 DR Interpro; IPR008957; TSP_N.
 DR Interpro; IPR003129; TSP_N.
 DR Interpro; IPR002035; VWF_A.
 DR Pfam; PF01391; Collagen; 5.
 DR Pfam; PF00041; fn3; 18.
 DR Pfam; PF02210; TSP_N; 1.
 DR Pfam; PF00092; VMA_4.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00060; FN3; 18.
 DR SMART; SM00210; TSPN; 1.
 DR SMART; SM00327; VMA; 4.
 DR PROSITE; PSS0853; FN3; 18.
 DR PROSITE; PSS0234; VMPA; 4.
 KW Alternative splicing; Cell adhesion; Collagen; Extracellular matrix; Glycoprotein; Hydroxylation; Repeat; Signal; Structural protein.
 FT SIGNAL 1 24
 FT CHAIN 25 3119
 FT DOMAIN 25 112 Collagen alpha 1(XII) chain.
 FT DOMAIN 140 316 Fibronectin type-III 1.
 FT DOMAIN 333 422 Fibronectin type-III 2.
 FT DOMAIN 444 620 VMPA 2.
 FT DOMAIN 635 723 Fibronectin type-III 3.
 FT DOMAIN 726 814 Fibronectin type-III 4.
 FT DOMAIN 817 905 Fibronectin type-III 5.
 FT DOMAIN 908 997 Fibronectin type-III 6.
 FT DOMAIN 999 1087 Fibronectin type-III 7.
 FT DOMAIN 1090 1179 Fibronectin type-III 8.
 FT DOMAIN 1203 1375 VMPA 3.

	FT	DOMAIN	1388	1476	Fibronectin type-III 9.
	FT	DOMAIN	1478	1567	Fibronectin type-III 10.
	FT	DOMAIN	1569	1656	Fibronectin type-III 11.
	FT	DOMAIN	1660	1747	Fibronectin type-III 12.
	FT	DOMAIN	1758	1847	Fibronectin type-III 13.
	FT	DOMAIN	1849	1937	Fibronectin type-III 14.
	FT	DOMAIN	1939	2028	Fibronectin type-III 15.
	FT	DOMAIN	2030	2119	Fibronectin type-III 16.
	FT	DOMAIN	2121	2208	Fibronectin type-III 17.
	FT	DOMAIN	2212	2298	Fibronectin type-III 18.
	FT	DOMAIN	2329	2501	WFA 4.
	FT	DOMAIN	2525	2717	TSP N-terminal.
	FT	DOMAIN	2456	2751	Nonhelical region (NC3).
	FT	DOMAIN	2752	2899	Triple-helical region (COL2) with 1 imperfection.
	FT	DOMAIN	2900	2942	Nonhelical region (NC2).
	FT	DOMAIN	2943	3045	Triple-helical region (COL1) with 2 imperfections.
	FT	DOMAIN	3046	3119	Nonhelical region (NC1).
	FT	SITE	866	868	Cell attachment site (Potential).
	FT	SITE	2784	2786	Cell attachment site (Potential).
	FT	SITE	2896	2898	Cell attachment site (Potential).
	FT	MOD_RES	2945	2945	Hydroxyproline (By similarity).
	FT	MOD_RES	2948	2948	Hydroxyproline (By similarity).
	FT	MOD_RES	2951	2951	Hydroxyproline (By similarity).
	FT	MOD_RES	2960	2960	Hydroxyproline (By similarity).
	FT	MOD_RES	2966	2966	Hydroxyproline (By similarity).
	FT	MOD_RES	2969	2969	Hydroxyproline (By similarity).
	FT	MOD_RES	2972	2972	Hydroxyproline (By similarity).
	FT	MOD_RES	2984	2984	Hydroxyproline (By similarity).
	FT	MOD_RES	3001	3001	Hydroxyproline (By similarity).
	FT	MOD_RES	3004	3004	Hydroxyproline (By similarity).
	FT	MOD_RES	3015	3015	Hydroxyproline (By similarity).
	FT	MOD_RES	3024	3024	Hydroxyproline (By similarity).
	FT	MOD_RES	3027	3027	Hydroxyproline (By similarity).
	FT	MOD_RES	3030	3030	Hydroxyproline (By similarity).
	FT	DOMAIN	869	872	Poly-Thr.
	FT	CARBOHYD	704	704	N-linked (GlcNAc . .) (Potential).
	FT	CARBOHYD	802	802	O-linked (XyI . . .) (Chondroitin sulfate)
	PT	CARBOHYD	893	893	(Potential).
	PT	CARBOHYD	985	985	O-linked (XyI . . .) (Chondroitin sulfate)
	FT	CARBOHYD	1769	1769	O-linked (XyI . . .) (Chondroitin sulfate)
	FT	CARBOHYD	2212	2212	(Potential).
	FT	CARBOHYD	2533	2533	N-linked (GlcNAc . . .) (Potential).
	FT	CARBOHYD	2684	2684	N-linked (GlcNAc . . .) (Potential).
	FT	VARSPLIC	25	1190	N-linked (GlcNAc . . .) (Potential).
	FT	VARSPLIC	3062	3064	Misling (In isoform XIIB-1 and isoform XIIB-2).
	FT	VARSPLIC	3119	3119	/FTID=VSP_001150.
	FT	VARSPLIC	3119	3119	Bpy -> GSG (In isoform XIIA-2 and isoform XIIB-2).
	FT	VARSPLIC	3119	3119	/FTID=VSP_001151.
	FT	VARSPLIC	3119	3119	Misling (In isoform XIIA-2 and isoform XIIB-2).
	FT	VARSPLIC	3119	3119	/FTID=VSP_001152.
	FT	VARSPLIC	3119	3119	Misling (In isoform XIIA-2 and isoform XIIB-2).
	FT	VARSPLIC	3119	3119	/FTID=VSP_001153.
	FT	VARSPLIC	3119	3119	Misling (In isoform XIIA-2 and isoform XIIB-2).
	FT	VARSPLIC	3119	3119	/FTID=VSP_001154.
	FT	VARSPLIC	3119	3119	Misling (In isoform XIIA-2 and isoform XIIB-2).
	FT	VARSPLIC	3119	3119	/FTID=VSP_001155.
	FT	VARSPLIC	3119	3119	Misling (In isoform XIIA-2 and isoform XIIB-2).
	FT	VARSPLIC	3119	3119	/FTID=VSP_001156.
	FT	VARSPLIC	3119	3119	Misling (In isoform XIIA-2 and isoform XIIB-2).
	FT	VARSPLIC	3119	3119	/FTID=VSP_001157.
	FT	VARSPLIC	3119	3119	Misling (In isoform XIIA-2 and isoform XIIB-2).
	FT	VARSPLIC	3119	3119	/FTID=VSP_001158.
	FT	VARSPLIC	3119	3119	Misling (In isoform XIIA-2 and isoform XIIB-2).
	FT	VARSPLIC	3119	3119	/FTID=VSP_001159.
	FT	VARSPLIC	3119	3119	Misling (In isoform XIIA-2 and isoform XIIB-2).
	FT	VARSPLIC	3119	3119	/FTID=VSP_001160.
	FT	VARSPLIC	3119	3119	Misling (In isoform XIIA-2 and isoform XIIB-2).
	FT	VARSPLIC	3119	3119	/FTID=VSP_001161.
	FT	VARSPLIC	3119	3119	Misling (In isoform XIIA-2 and isoform XIIB-2).
	FT	VARSPLIC	3119	3119	/FTID=VSP_001162.
	FT	VARSPLIC	3119	3119	Misling (In isoform XIIA-2 and isoform XIIB-2).
	FT	VARSPLIC	3119	3119	/FTID=VSP_001163.
	FT	VARSPLIC	3119	3119	Misling (In isoform XIIA

Db	557	SSDAF----	RDBAIKLRNSDVEIFAVGVADVA	SELEAIIASPPAE	THVFTVED	FD	AFQRI	61
RESULT 11								
Q8T6U5		PRELIMINARY;	PRT;	441	AA.			
ID	Q8T6U5							
AC	Q8T6U5;							
DT	01-JUN-2002 (TREMBLrel. 21, Created)							
DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)							
DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)							
DE	Proximal thread matrix protein 1 variant a.							
OS	Mytilus edulis (Blue mussel).							
CC	Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;							
OC	Mytiloidea; Mytilidae; Mytilus.							
OX	NCBI_TaxID=6550;							
RN	[1]							
RP	SEQUENCE FROM N.A.							
RX	MDLINB=22313036; PubMed=12425661; DOI=10.1021/bm0255903;							
RA	Sun C., Lucac J.M., Walte J.H.;							
RT	"Collagen-binding matrix proteins from elastomeric extraorganic							
RL	bysal fibers.";							
DR	Biocomolecules 3:1240-1248(2002).							
DR	EMBL; AF41454; AAL63537.1; -							
DR	GO; GO:0005198; F:structural molecule activity; IEA.							
DR	InterPro: IPR002035; VMF_A.							
DR	Pfam; PF00092; VMA, 2.							
DR	PRINTS; PR0453; VMFADOMAIN.							
DR	SMART; SM00327; VMA; 2.							
DR	PROSITE; PS50234; VMA; 2.							
KW	Matrix protein.							
SQ	SEQUENCE 441 AA; 47543 MW; 881DBBD36B891D2B CRC64;							
Query Match		13.9%;	Score 125;	DB 2;	Length 441;			
Best Local Similarity		24.7%;	Pred. No. 0.016;					
Matches	47;	Conservative	44;	Mismatches	73;	Indels	26;	Gaps 12
Qy	1	DLYPLDLSGGV----	LHNMIEIYFVYQLAHKF--	ISPO-LRMSFIYFSTGTTLMKLT	53			
Db	241	DIATVFDSSISINANNPNNYGLMKDFMKDIDRFNKTGPDGTQFAVVTFADPATKQFGLK	300					
Qy	54	E--DREOIRQGLEELQKVLPG--	GDYVHSGFERASEQIYYENRQSG--	YTPASVIALTL	106			
Db	301	DYSSKAEIKGKID---KTPSIIIGTATAGDGLENARLEV--	FPNRNGGGRGEVQKVILLT	356				
Qy	107	DGEL--HEDLFYSEERENRSDIGALVYCVGV--	KDFNETQLARIADSKDHVPVNDGQ	163				
Db	357	DGQNNHGHS----DEHESLLRKGGVIVAI	IGVGTGFPKSEILINIASSEYVF--	TTSSFN	411			
Qy	164	ALQGIHSIL	173					
Db	412	KLSKIMENVV	421					
RESULT 12								
Q8T5C3		PRELIMINARY;	PRT;	444	AA.			
ID	Q8T5C3							
AC	Q8T5C3;							
DT	01-JUN-2002 (TREMBLrel. 21, Created)							
DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)							
DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)							
DE	Proximal thread matrix protein 1b.							
OS	Mytilus edulis (Blue mussel).							
CC	Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;							
OC	Mytiloidea; Mytilidae; Mytilus.							
OX	NCBI_TaxID=6550;							
RN	[1]							
RP	SEQUENCE FROM N.A.							
RX	MDLINB=22313036; PubMed=12425661; DOI=10.1021/bm0255903;							
RA	Sun C., Lucac J.M., Walte J.H.;							
RT	"Collagen-binding matrix proteins from elastomeric extraorganic							
RL	bysal fibers.";							
DR	Biocomolecules 3:1240-1248(2002).							
DR	EMBL; AY053390; AAL17973.1; -							

DR GO:0005198; Functional molecule activity; IEA.
 DR InterPro: IPR002035; VMA_A.
 DR Pfam: PF00092; VMA; 2.
 DR PRINTS: PR00453; VMAFADOMAIN.
 DR SMART: SM00327; VMA; 2.
 DR PROSITE: PSS0234; VMA; 2.
 KM Matrix protein.
 SQ SEQUENCE 444 AA; 4781 MW; D2C605347450C931 CRC64;
 Query Match 13.9%; Score 125; DB 2; Length 444;
 Best local similarity 24.7%; Pred. No 0.016;
 Matches 47; Conservative 44; Mismatches 73; Indels 26; Gaps 12;
 QY 1 DLYFIIDKSGSV---LHMNEIYFVEQLAHKF--ISPO-LRMSFVFTSTRTTLTKLT 53
 DB 244 DIAFVFPASSISNNNNNNYGLMKDFMKDIYDRNKTGPDGTGPAVATFADRAIKGGLK 303
 QY 54 E--DREIDRGLELQVVG--GDYMHGSPSPASQIYENQ--YRTASVITALT 106
 DB 304 DYSSKAEIKGAID---KVTPTSIGQTAIGDLNARLEV-FPNNMGGRREVCKVILLT 359
 QY 107 DGEI--HEDLFPYSEKANSRDLGATVCGV-KDFTNQLARIADSKDHPVNDGFG 163
 DB 360 DQGNHKS---PEHSSLRKEGVVAIVGTGTGLKSLINIASSEIVF-TTSSFN 414
 QY 164 ALQGIHSIL 173
 DB 415 KLSKIMENVV 424
 RESULT 13
 CALC CHICK STANDARD; PRT, 3124 AA.
 ID CA1C CHICK 004509;
 AC P13944; 004509;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Collagen alpha 1(XII) chain precursor (Fibrochimerin).
 GN Name=COL12A1;
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=White Leghorn;
 RX Yamaoka M., Yamada K.M., Yamada S.S., Shinomura T., Tanaka H.,
 RA Nishida Y., Ohara M., Kimata K.;
 RT "The complete primary structure of type XII collagen shows a chimeric
 RT molecule with reiterated fibronectin type III motifs, von Willebrand
 RT factor A motifs, a domain homologous to a noncollagenous region of
 RT type IX collagen, and short collagenous domains with an Arg-Gly-Asp
 RT site."
 RL J. Cell Biol. 115:209-221(1991).
 RN [2]
 RP SEQUENCE OF 2456-3124 FROM N.A., AND SEQUENCE OF 2772-2794 AND
 RP 2846-2873.
 RX MEDLINE=90062079; PubMed=2584192;
 RA Gordon M.K., Gerecke D.R., Dubler B., van der Reet M., Olsen B.R.;
 RT "Type XII collagen. A large multidomain molecule with partial homology
 RT to type IX collagen."
 RL J. Biol. Chem. 264:19772-19778(1989).
 RN [3]
 RP SEQUENCE OF 2960-3076 FROM N.A.
 RX MEDLINE=97317590; PubMed=476925;
 RA Gordon M.K., Gerecke D.R., Olsen B.R.;
 RT "Type XII collagen: distinct extracellular matrix component discovered
 RT by cDNA cloning."
 RL Proc. Natl. Acad. Sci. U.S.A. 84:6040-6044(1987).
 RN [4]
 RP SEQUENCE OF 1-1283 FROM N.A. (ISOFORM SHORT), AND ALTERNATIVE

RP SPLICING.
 RC TISSUE=Embryo;
 RX MEDLINE=93042014; PubMed=1420368; DOI=10.1016/0167-4781(92)90145-P;
 RA Trueb J., Trueb B.;
 RT "The two splice variants of collagen XII share a common 5' end."
 RL Biochim. Biophys. Acta 1171:97-98(1992).
 RN [5]
 RP ALTERNATIVE SPLICING.
 RX MEDLINE=95370352; PubMed=7642694; DOI=10.1083/jcb.130.4.1005;
 RA Koch M., Bohmann B., Matchison M., Hagios C., Trueb B., Chiquet M.;
 RT "Large and small splice variants of collagen XII: differential
 RT expression and ligand binding."
 RL J. Cell Biol. 130:1005-1014(1995).
 CC -1- FUNCTION: Type XII collagen interacts with type I collagen-
 CC containing fibrils, the COL1 domain could be associated with the
 CC surface of the fibrils, and the COL2 and NC3 domains may be
 CC localized in the pericellular matrix.
 CC -1- SUBUNIT: Trimer of identical chains each containing 190 kDa of
 CC nontriple-helical sequences.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Comment=The final tissue form of collagen XII may contain
 CC homotrimers of either isoform long or isoform short or any
 CC combination of isoform long and isoform short. Only isoform long
 CC is a proteoglycan. Isoform long has more restricted expression
 CC in embryonic tissue than isoform short;
 CC Name=long;
 CC IsoId=P13944-1; Sequence=Displayed;
 CC Name=short;
 CC IsoId=P13944-2; Sequence=VSP_001148;
 CC -1- TISSUE SPECIFICITY: Type XII collagen is present in tendons,
 CC ligaments, perichondrium, and peritoneum, all dense connective
 CC tissues containing type I collagen.
 CC -1- DOMAIN: This sequence defines five distinct domains, two triple-
 CC helical domains (COL1 and COL2) and three nontriple-helical
 CC domains (NC1, NC2, and NC3).
 CC -1- PTM: The triple-helical tail is stabilized by disulfide bonds at
 CC each end.
 CC -1- PTM: Prolines at the third position of the tripeptide repeating
 CC unit (G-X-Y) are hydroxylated in some or all of the chains.
 CC -1- PTM: O-glycosylated; glycosaminoglycan of chondroitin-sulfate type
 CC (by similarity).
 CC -1- SIMILARITY: Belongs to the fibril-associated collagens with
 CC interrupted helices (FACIT) family.
 CC -1- SIMILARITY: Contains 18 fibronectin type III domains.
 CC -1- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
 CC -1- SIMILARITY: Contains 4 VMA domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D00824; BAA00701.1; -;
 DR EMBL: X61024; CAA43358.1; -;
 DR EMBL: M17375; AAA48718.1; -;
 DR EMBL: J05137; AAA48635.1; -;
 DR EMBL: X67327; CAA47744.1; -;
 DR PIR: A40020; A40020.
 DR HSP: P56199; 10C5.
 DR InterPro: IPR008160; Collagen.
 DR InterPro: IPR008985; Cona_like_1ec_g1.
 DR InterPro: IPR003961; FN III.
 DR InterPro: IPR008957; FN III-like.
 DR InterPro: IPR003129; TSP N.
 DR InterPro: IPR002035; VMA_A.
 DR Pfam: PF01391; Collagen; 4.
 DR Pfam: PF00041; Fn3; 17.
 DR Pfam: PF02210; TSP N; 1.
 DR Pfam: PF00092; VMA; 4.

```

DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00060; FN3; 18.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00327; VMA; 4.
DR PROSITE; PS50853; FN3; 18.
DR PROSITE; PS50234; VMA; 4.
DR Alternative splicing; Cell adhesion; Collagen;
KM Direct protein sequencing; Extracellular matrix; Glycoprotein;
KW Hydroxylation; Repeat; Signal; Structural protein.
FT SIGNAL 1 24
FT CHAIN 25 3124
FT DOMAIN 25 112
FT DOMAIN 139 311
FT DOMAIN 332 421
FT DOMAIN 439 615
FT DOMAIN 630 718
FT DOMAIN 721 809
FT DOMAIN 812 902
FT DOMAIN 905 993
FT DOMAIN 995 1083
FT DOMAIN 1086 1175
FT DOMAIN 1199 1371
FT DOMAIN 1386 1472
FT DOMAIN 1474 1564
FT DOMAIN 1566 1654
FT DOMAIN 1655 1745
FT DOMAIN 1756 1845
FT DOMAIN 1847 1935
FT DOMAIN 1937 2026
FT DOMAIN 2028 2117
FT DOMAIN 2119 2206
FT DOMAIN 2210 2294
FT DOMAIN 2327 2500
FT DOMAIN 2524 2716
FT DOMAIN 2455 2750
FT DOMAIN 2751 2902
FT DOMAIN 2903 2945
FT DOMAIN 2946 3048
FT DOMAIN 3049 3124
FT DOMAIN 3086 3096
FT DOMAIN 3111 3123
FT SITE 2899 2901
FT CARBOHYD 32 32
FT CARBOHYD 797 797
FT CARBOHYD 890 890
FT CARBOHYD 981 981
FT CARBOHYD 981 981
FT CARBOHYD 1006 1006
FT CARBOHYD 1032 1032
FT CARBOHYD 1044 1044
FT CARBOHYD 1512 1512
FT CARBOHYD 1767 1767
FT CARBOHYD 2210 2210
FT CARBOHYD 2273 2273
FT CARBOHYD 2532 2532
FT CARBOHYD 2683 2683
FT VARSPLIC 25 1188
FT CONFLICT 1258 1258
FT CONFLICT 1264 1264
FT CONFLICT 2759 2759
FT CONFLICT 2803 2803
FT CONFLICT 2977 2977
FT CONFLICT 3075 3075
SQ SEQUENCE 3124 AA; 340578 MW; 094285AF7F346CF CRC64;

Query Match 13.9%; Score 125; DB 1; Length 3124;
Best Local Similarity 27.1%; Pred. No. 0.15;
Matches 48; Conservative 39; Mismatches 16; Gaps 9;

```

```

QY 1 DLVFIIDKSGSV-LHMHNEIYYVEQLAKF--ISPO-LMMSFVSTRGTTLMKL--E 54
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 2327 DIVFLTDASMSIDBDNFNKVKKEVFNTVGAFDLINFAGIQVSVQVSDEAQSEKPLNTFPD 2386
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 55 DREGIRQGLEELQKVLPGGDTVMHGEFPERASEOIV-YENNGYRTASVIALTDGELHED 113
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 2387 DKQALGALQNVQ--YRGGTRTGKALTFTIKKVLTVESGMRGVPKVLVAVTDGNSODE 2444
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 114 LFPPYSREANRSDLAIVYCVGKDFNETQLARIAD--SKDHVPVNDGFOALOGI 168
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 2445 V---RKAATVIGHSGPSVFGVADVVDYVELAKISKPSRRVFIYVD--FDAPEKI 2496
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
Q9UBJ9 PRELIMINARY; PRT; 765 AA.
AC Q9UBJ9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Thrombospondin-related adhesive protein homolog.
GN Name=MIC2;
OS Neospora caninum.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Sarcocystidae;
OC Neospora.
OX NCBI_TaxID=29176;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20183852; PubMed=10717300; DOI=10.1016/S0166-6851(99)00228-5;
RA Lovett J.L., Howe D.K., Sibley L.D.;
RL "Molecular characterization of a thrombospondin-related anonymous
protein homologue in Neospora caninum."
EMBL AF061273; AAF01565.1; -.
DR HSSP; P07996; 1LSL.
DR GO; GO:0008083; P: growth factor activity; IEA.
DR InterPro; IPR001969; Rept_Asp_AS.
DR InterPro; IPR00762; PTN_MK.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; TSP_1; 6.
DR Pfam; PF00092; VMA_1.
DR PRINTS; PRO0453; VWFADOMAIN.
DR SMART; SM00209; TSP1; 6.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.
DR PROSITE; PS50092; TSP1; 2.
DR PROSITE; PS50234; VMA; 1.
SQ SEQUENCE 765 AA; 82880 MW; 9727838C1FE8CD4 CRC64;

Query Match 13.8%; Score 124; DB 2; Length 765;
Best Local Similarity 27.5%; Pred. No. 0.036;
Matches 41; Conservative 27; Mismatches 61; Indels 20; Gaps 7;

QY 1 DLVFIIDKSGSV-LHMHNEIYYVEQLAKFIS-----POLMSFVSTRGTTLMKL- 52
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 77 DICEFLVDSGSGIGEAHYEE---VKQFLHAFUSKLPDIGNDEVNTSLVISTYVHPHMSLR 132
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 53 ---TEDREGIRQGLEELQKVLPGGDTVMHGEFPERASEOIV-YENNGYRTASVIALTDG 108
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 133 ANNASKEKRAMQ--DVLTIFPHGGTTNTAAGLOTCTQMPLFDYRERQVPPKVLVMTDGS 190
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 109 ELHEDLFYSERANRSDLAIVYCVG 137
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 191 ESDSD--FHTVNAKVIIRRGGITVLSV 217
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
ITAD HUMAN
ID ITAD HUMAN STANDARD; PRT; 1162 AA.
AC Q13349; Q15575; Q15576;
DT 16-OCT-2001 (Rel. 40, Created)

```

DT 16-OCT-2001 (Rel. 40, last sequence update)
DE Integrin alpha-D precursor (leukointegrin alpha D) (CD11d) (ADB2).
GN Name:ITGAD;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=96111956; PubMed=8777714; DOI=10.1016/1074-7613(95)90058-6;
RA Van der Vlieten M., Le Trong H., Wood C.L., Moore P.F., St John T.,
RA Staunton D.E., Gallatin W.M.;
RT "A novel leukointegrin, alpha d beta 2, binds preferentially to ICAM-3.";
RL Immunity 3:683-690(1995).
RN [2]
RP SEQUENCE OF 1-235 FROM N.A.
RX MEDLINE=20187620; PubMed=10722744; DOI=10.1074/jbc.275.12.8959;
RA Noci J.D., Johnson A.K., Dillon J.D.;
RT "Structural and functional characterization of the leukocyte integrin gene CD1d. Essential role of Sps and Sp3.";
RL J Biol. Chem. 275:8959-8969(2000).
RN [3]
RP SEQUENCE OF 467-537; 571-602; 633-666; 788-834 AND 910-1125 FROM N.A.
RX MEDLINE=96257236; PubMed=8666289; DOI=10.1016/0378-1119(95)00869-1;
RA Wong D.A., Davis E.M., LeBeau M., Springer T.A.;
RT "Cloning and chromosomal localization of a novel gene-encoding a human beta 2-integrin alpha subunit.";
RL Gene 171:291-294(1996).
RN [4]
RP INTERACTION WITH VCAM1.
RX MEDLINE=99059842; PubMed=9841932;
RA Grayson M.H., Van der Vlieten M., Sterbinsky S.A., Michael Gallatin W.,
RA Hoffman P.A., Staunton D.E., Bochner B.S.;
RT "Alphadelta2 integrin is expressed on human eosinophils and functions as an alternative ligand for vascular cell adhesion molecule 1 (VCAM-1).";
RL J. Exp. Med. 188:2187-2191(1998).
RN [5]
RP INTERACTION WITH VCAM1.
RX MEDLINE=99370002; PubMed=10438935;
RA Van der Vlieten M., Crowe D.T., Hoekstra D., Vazquez R., Hoffman P.A.,
RA Grayson M.H., Bochner B.S., Gallatin W.M., Staunton D.E.;
RT "The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a binding interface between I domain and VCAM-1.";
RL J. Immunol. 163:1984-1990(1999).
CC -1- FUNCTION: Integrin alpha-D/beta-2 is a receptor for ICAM3 and VCAM1. May play a role in the atherosclerotic process such as clearing lipoproteins from plaques and in phagocytosis of blood-borne pathogens, particulate matter, and senescent erythrocytes from the blood.
CC -1- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-D associates with beta-2.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: Expressed moderately on myelomonocytic cell lines and subsets of peripheral blood leukocytes and strongly on tissue-specialized cells, including macrophages foam cells within atherosclerotic plaques, and on splenic red pulp macrophages.
CC -1- DOMAIN: The integrin I-domain (insert) is a VWA domain. Integrins with I-domains do not undergo protease cleavage.
CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
CC -1- SIMILARITY: Contains 1 VWA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----
DR EMBL; U37028; AAB38547.1; -;
DR EMBL; U40274; AAB60634.1; -;
DR EMBL; U40275; AAB60635.1; -;
DR EMBL; U40276; AAB60636.1; -;
DR EMBL; U40277; AAB60637.1; -;
DR EMBL; U40279; AAB60638.1; -;
DR EMBL; U40278; AAB60638.1; JOINED.
DR EMBL; AF187881; AAF62875.1; -;
DR HSSP; P11215; IBHQ.
DR Gene; HGNC:6146; ITGAD.
DR MIM; 602453; -;
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0016337; P:cell-cell adhesion; NAS.
DR GO; GO:0007160; P:cell-matrix adhesion; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR02035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_alpha; 1.
DR Pfam; PF00092; VWA; 1.
DR PRINTS; PRO1185; INTEGRINA.
DR PRINTS; PRO0453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWA; 1.
KW Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
KW Repeat; Signal; Transmembrane
FT SIGNAL 1 17 Potential.
FT CHAIN 18 1162 Integrin alpha-D.
FT DOMAIN 18 1100 Extracellular (Potential).
FT TRANSMEM 1101 1121 Potential.
FT DOMAIN 1122 1162 Cytoplasmic (Potential).
FT REPEAT 32 85 FG-GAP 1.
FT REPEAT 86 97 FG-GAP 2.
FT REPEAT 150 332 VWF_A.
FT REPEAT 350 400 FG-GAP 3.
FT REPEAT 401 452 FG-GAP 4.
FT REPEAT 454 516 FG-GAP 5.
FT REPEAT 518 576 FG-GAP 6.
FT REPEAT 581 633 FG-GAP 7.
FT CA_BIND 465 473 Potential.
FT CA_BIND 530 538 Potential.
FT CA_BIND 593 601 Potential.
FT SITE 1127 1131 GPCR motif.
FT DISULFID 67 74 By similarity.
FT DISULFID 106 124 By similarity.
FT DISULFID 655 710 By similarity.
FT DISULFID 769 775 By similarity.
FT DISULFID 846 861 By similarity.
FT DISULFID 994 1018 By similarity.
FT DISULFID 1023 1028 By similarity.
FT CARBOHYD 59 59 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 87 87 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 99 99 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 391 391 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 691 691 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 733 733 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 873 873 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 957 957 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1046 1046 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1046 1046 N-linked (GlcNAc...) (Potential).
FT CONFLICT 500 500 Missing (in Ref. 2).
FT CONFLICT 515 518 GHFV -> ATP (in Ref. 2).
FT CONFLICT 825 825 L -> V (in Ref. 2).
FT CONFLICT 984 984 V -> A (in Ref. 2).
SQ SEQUENCE 1162 AA; 126885 MW; F236A1A3545D77D CRC64;
Query March 13.6%; Score 122.5; DB 1; Length 1162;
Best Local Similarity 26.0%; Pred. No. 0.08;
Matches 47; Conservative 35; Mismatches 80; Indels 19; Gaps 7;
Qy 1 DLYEILDKSGSV-LHHNNEIYFVEQLAHKRFISPOLRMSPFVSTRTGTLMLKTEDRRQI 59

```

Db      150  DIYFLIDSGSSIDQNDPNOCKGPFQAVMGQPEGTDTLPALMOYS---NLKIHFTFTOF 205
QY      60  RQGLEELQKVL P---GADTYMHEGFPERASEQIYYENRQGYRTA-SVYIALTDGELHEDL 114
Db      206  RTSBPSQOSLVDPIVOLKGLFTATGILTVVTLPHHNGARKSAKILIVITDQKYKDP 265
QY      115  FYSEREARNSRDLGALVYCVGVKDFNETQLAR-----IADSKDHVPFVNDGFQALOG 167
Db      266  LEYSD-VIPOAEKAGIIRYAIGVGHAFQGPTRQELNTISSAPPQDHVFKV-DNFALGS 323
QY      168  I 168
Db      324  I 324

```

Search completed: June 13, 2005, 20:03:16
 Job time : 59.0753 secs